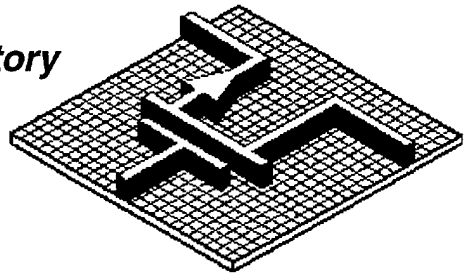
Hardware Technologies for Robust Personal Communication Transceivers

FINAL TECHNICAL REPORT

Contract No. DAAB07-93-C-C501
Performance Period: July 1992 -- July 1995
Program Manager: Dr. Ken Gabriel

Henry Samueli
Asad A. Abidi
Gregory J. Pottie
Yahya Rahmat-Samii

Integrated Circuits & Systems Laboratory
Electrical Engineering Department
University of California
Los Angeles, CA 95-1594



Prepared for

Defense Advanced Research Projects Agency
Electronic Technology Office

DISCLAIMER NOTICE

THIS DOCUMENT IS BEST QUALITY AVAILABLE. THE COPY FURNISHED TO DTIC CONTAINED A SIGNIFICANT NUMBER OF PAGES WHICH DO NOT REPRODUCE LEGIBLY.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	July 1995	Final Report: 7/7/92 - 7/7/95

4. TITLE AND SUBTITLE
"Hardware Technology for Robust Personal Communication Transceivers: Final Report"

5. FUNDING NUMBERS
DAAB07-93-C-C501

6. AUTHOR(S)
Henry Samueli, P.I. Yahya Rahmat-Samii, Co-P.I.
Asad Abidi, Co-P.I.
Greg Pottie, Co-P.I.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
UCLA - Electrical Engineering Dept.
56-125B, Engineering IV, Box 951594
Los Angeles, CA 90095-1594
(310) 825-6783

8. PERFORMING ORGANIZATION REPORT NUMBER
N/A

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Commander US Army CECOM
Amsel-ACCA-D-AW
Ft. Monmouth, NJ 07703-5008
(908) 532-1530

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
Not Known

11. SUPPLEMENTARY NOTES
N/A

12a. DISTRIBUTION/AVAILABILITY STATEMENT
APPROVED FOR PUBLIC RELEASE; DISTRIBUTION IS UNLIMITED.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

The overall goal of this research project is to develop low-power personal communications transceiver hardware technologies, coupled with advanced systems techniques such as antenna diversity, channel coding, and adaptive power control, for achieving robust wireless digital data transmission over multipath fading channels. A frequency-hopped, spread-spectrum (FH/SS), code-division, multi-le-access (CDMA) technique was chosen over other multiple access schemes because: a) it provides an inherent immunity to multipath fading; and b) signal processing is performed at the hopping rate, which is much slower than the chip rate encountered either in a direct-sequence (DS) CDMA or time-division, multi-le access (TDMA) system, thereby potentially resulting in much lower receiver power consumption. The project is organized into four major tasks:
1) System Architecture/Performance Evaluation (Prof. G. Pottie)
2) Receiver Architecture/Signal Processing (Prof. H. Samueli)
3) Analog Systems/Circuits (Prof. A. Abidi)
4) Miniature Antenna Design (Prof. Y. Rahmat-Samii)

14. SUBJECT TERMS
spread-spectrum, wireless transceivers, frequency-hopping, CMOS RF Circuits.

15. NUMBER OF PAGES
294

16. PRICE CODE
N/A

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	

GENERAL INSTRUCTIONS FOR COMPLETING SF 298

The Report Documentation Page (RDP) is used in announcing and cataloging reports. It is important that this information be consistent with the rest of the report, particularly the cover and title page. Instructions for filling in each block of the form follow. It is important to ***stay within the lines*** to meet ***optical scanning requirements***.

Block 1. Agency Use Only *(Leave blank)*.

Block 2. Report Date. Full publication date including day, month, and year, if available (e.g. 1 Jan 88). Must cite at least the year.

Block 3. Type of Report and Dates Covered. State whether report is interim, final, etc. If applicable, enter inclusive report dates (e.g. 10 Jun 87 - 30 Jun 88).

Block 4. Title and Subtitle. A title is taken from the part of the report that provides the most meaningful and complete information. When a report is prepared in more than one volume, repeat the primary title, add volume number, and include subtitle for the specific volume. On classified documents enter the title classification in parentheses.

Block 5. Funding Numbers. To include contract and grant numbers; may include program element number(s), project number(s), task number(s), and work unit number(s). Use the following labels:

C	- Contract	**PR**	- Project
G	- Grant	**TA**	- Task
PE	- Program Element	**WU**	- Work Unit Accession No.

Block 6. Author(s). Name(s) of person(s) responsible for writing the report, performing the research, or credited with the content of the report. If editor or compiler, this should follow the name(s).

Block 7. Performing Organization Name(s) and Address(es). Self-explanatory.

Block 8. Performing Organization Report Number. Enter the unique alphanumeric report number(s) assigned by the organization performing the report.

Block 9. Sponsoring/Monitoring Agency Name(s) and Address(es). Self-explanatory.

Block 10. Sponsoring/Monitoring Agency Report Number. *(If known)*

Block 11. Supplementary Notes. Enter information not included elsewhere such as: Prepared in cooperation with...; Trans. of...; To be published in.... When a report is revised, include a statement whether the new report supersedes or supplements the older report.

Block 12a. Distribution/Availability Statement. Denotes public availability or limitations. Cite any availability to the public. Enter additional limitations or special markings in all capitals (e.g. NOFORN, REL, ITAR).

- **DOD** - See DoDD 5230.24, "Distribution Statements on Technical Documents."
- **DOE** - See authorities.
- **NASA** - See Handbook NHB 2200.2.
- **NTIS** - Leave blank.

Block 12b. Distribution Code.

- **DOD** - Leave blank.
- **DOE** - Enter DOE distribution categories from the Standard Distribution for Unclassified Scientific and Technical Reports.
- **NASA** - Leave blank.
- **NTIS** - Leave blank.

Block 13. Abstract. Include a brief *(Maximum 200 words)* factual summary of the most significant information contained in the report.

Block 14. Subject Terms. Keywords or phrases identifying major subjects in the report.

Block 15. Number of Pages. Enter the total number of pages.

Block 16. Price Code. Enter appropriate price code *(NTIS only)*.

Blocks 17. - 19. Security Classifications. Self-explanatory. Enter U.S. Security Classification in accordance with U.S. Security Regulations (i.e., UNCLASSIFIED). If form contains classified information, stamp classification on the top and bottom of the page.

Block 20. Limitation of Abstract. This block must be completed to assign a limitation to the abstract. Enter either UL (unlimited) or SAR (same as report). An entry in this block is necessary if the abstract is to be limited. If blank, the abstract is assumed to be unlimited.

TABLE OF CONTENTS

I. PROJECT GOALS 1

II. DETAILED STATUS REPORT 2

A. Systems 2
1. System Architecture and Rationale 2
2. Power Control and Coding Simulation 2
3. Synchronization and System Simulation 3

B. Architectures 4
1. RF/IF Architecture 4
2. Baseband Architecture 5

C. Advanced Analog and Digital Circuits 6
1. Spread-Spectrum RF to Baseband Receiver 6
2. Local Oscillator 7
3. Power Amplifier 8
4. Spread Spectrum to RF Transmitter 9
5. Digital Baseband Receiver 10

D. Miniature Antenna Design 11
1. Analysis and Design Techniques 11

III. PUBLICATIONS AND PRESENTATIONS 13

IV. PERSONNEL 16

V. APPENDIX: REPRINTS OF PUBLICATIONS 17

I. Project Goals

The overall goal of this research project is to develop low-power personal communications transceiver hardware technologies, coupled with advanced systems techniques such as antenna diversity, channel coding, and adaptive power control, for achieving robust wireless digital data transmission over multipath fading channels. A frequency-hopped, spread-spectrum (FH/SS), code-division, multiple-access (CDMA) technique was chosen over other multiple access schemes because: a) it provides an inherent immunity to multipath fading; and b) signal processing is performed at the hopping rate, which is much slower than the chip rate encountered either in a direct-sequence (DS) CDMA or time-division, multiple access (TDMA) system, thereby potentially resulting in much lower receiver power consumption. Furthermore, frequency-shift keying (FSK) modulation with noncoherent detection in a frequency-hopped system results in a much simpler transceiver architecture as compared with coherent amplitude and phase modulation methods commonly used in direct sequence and TDMA systems. Architectural innovations as well as advanced circuit techniques are incorporated into the design of a portable handset to achieve minimum power and size without sacrificing robustness in performance. The transceiver integrates all RF, IF, and baseband processing components into a monolithic CMOS solution. Diversity techniques are also integrated into the handset to achieve the maximum diversity benefit for robust data transmission. The hardware and system technologies developed for the transceiver design are general enough to be applicable to a wide variety of commercial and military wireless communications applications such as cellular and micro-cellular radios and telephones, wireless LANs, and wireless PBX systems. The project is organized into four major tasks:

- **System Architecture/Performance Evaluation** (Prof. G. Pottie)
- **Receiver Architecture/Signal Processing** (Prof. H. Samueli)
- **Analog Systems/Circuits** (Prof. A. Abidi)
- **Miniature Antenna Design** (Prof. Y. Rahmat-Samii)

In order to validate the design techniques being proposed, a prototype all-CMOS MHz transceiver IC was developed. The specifications of this FH/SS transceiver are summarized in Table 1. This is the most highly integrated monolithic radio transceiver IC reported to date, and represents a significant milestone in the development of CMOS RF technology.

Table 1. FH/SS Transceiver Specifications

Multiple Access Scheme	Frequency-Hopped CDMA
Modulation Scheme	Quaternary/Binary FSK
Duplexing Mode	Time Division Duplex (TDD)
Channel Coding	Convolutional Code (R=1/2, K=6)
Hopping Bandwidth	26 MHz (902 - 928 MHz)
Data Rates	2 - 160 kbit/s
Maximum Hop Rate	1 hop per 8 symbols
Antenna Diversity	2 with separate rx channels
Maximum Range	500 meters
Transmission Power	20 mW - 20 mW
Active Power	300 mW (Tx), 225 mW (Rx)
Operating Voltage	3 V
IC Technology	1.0-μm CMOS

II. Detailed Status Report

A. Systems

1) System Architecture and Rationale (Prof. G. Pottie)

The objective of this personal communications project is to pass data at rates up to 160 kb/s between two low-power handsets, using the 902 to 928 MHz band in a microcellular environment. The principal channel impairments in multi-user radio systems are: a) attenuation due to shadowing, b) multipath fading, and c) interference from other radios.

To overcome these impairments without resorting to high transmitter power, the architecture developed in this project incorporates many advanced techniques. Two antennas are used to provide polarization/ space diversity. Equal gain diversity combining is employed for the two receive branches, to meet the high hopping rate constraint and simplify the hardware requirements. Frequency hopping/Code Division Multiple Access (CDMA) combined with error-control coding has been implemented to provide both frequency and interferer diversity, reducing the effects of multipath fading and interference from other radios. Using a combination of diversity techniques provides an economical means of achieving a high aggregate diversity order, thereby taming the radio channel and allowing robust receiver performance close to that achievable in the additive white Gaussian noise channel. Distributed adaptive power control is also employed, so that the minimum transmitted power required for reliable communication is used in order to prolong battery life as well as increase network capacity. Distributed power control with respect to a signal-to-interference threshold does not require central control and is well suited to peer-to-peer networks. Finally, an adaptive data rate permits a longer transmission range for lower rate messages without increasing the peak transmit power.

Frequency hopped systems permit a large frequency/interferer diversity to be obtained with lower transmission delay and significantly less complexity than wideband, direct sequence, spread-spectrum or time-division, multiple-access systems. Conversely, fast hopping generally precludes the use of a coherent receiver, implying a significant SNR penalty. However, the combination of frequency diversity, synchronous access and reduced complexity, compared to DS/SS CDMA, more than makes up for this penalty.

2) Power Control and Coding Simulation (V. Lin, Prof. G. Pottie)

A distributed power control scheme has been devised to meet the unique constraints of the frequency-hopped, spread-spectrum transceiver, and a multi-user simulation of power control for a peer-to-peer communication network of transceivers has been developed. The power control scheme is an augmentation of the work performed for an earlier DARPA research project. UCLA researchers, Chen et al., have formulated a distributed power control scheme based on a closed loop algorithm which updates the transmit power with measurements of signal-to-interference ratio (SIR) at the receiver. Their scheme includes a robust power update algorithm that protects the operational links by suppressing (blocking) new links when the quality of service for existing links falls below an acceptable threshold. We have modified some aspects of Chen's scheme so that it can be implemented within the constraints of the frequency-hopped system.

The capacity of clustered peer-to-peer FH/CDMA networks was computed by simulation. A simple algorithm was used to organize links into clusters. Each cluster has the same number of users. The simulation of a single-cluster network with our distributed algorithm shows that the slot-efficiency is approximately 6% at 1% outage probability when the SNR requirement is 9 dB. The outages can be caused by blocked as well as dropped calls. This performance figure is very poor in comparison to the

62% slot-efficiency for the single-cell network, which contains a base-station. The large disparity in performance between the clustered and cellular systems indicates a fundamental trade-off between capacity and network infrastructure. Although the only source of interference is from adjacent channel interference, ACI could cause severe problems when the links are randomly co-located.

One way to obtain reasonable capacity for peer-to-peer systems is to add more complexity by better frequency planning. For example, the guard band between different frequency slots can be increased to reduce ACI. Suppose a larger guard band is introduced so that there are fewer channels in the available bandwidth and that the ACI is reduced to -50 dB, our simulation shows that a slot-efficiency of about $16/32 = 50\%$ can be achieved.

From the network customer's perspective, dropped calls which interrupt on-going conversations are extremely undesirable. One solution is to employ a distributed admission control policy that allows voluntary termination of admission attempts when an operational link is in danger of being dropped. A local mechanism to sense the "resistance" of network due to other users is required for implementing a distributed policy. Since more calls are dropped at higher congestion levels, blocking or terminating admission attempts that encountered high resistance would effect a reduction in call dropping probability with a slight increase in blocking probability.

For peer-to-peer networks, simulation experiments show that blocking new links which see no SIR improvements despite increasing power can significantly reduce dropped calls, which will be referred to as Algorithm I. Algorithm II refines Algorithm I by testing both initial received power and gross power increase against corresponding thresholds to further reduce the call dropping probability. The simulation results for the single cluster peer-to-peer network with different admission control algorithms are summarized in Table 2. The dynamic range for the single-cluster system was set to 60 dB because the

Table 2. Call blocking and call dropping probabilities with and without distributed admission control. Single-Cluster Network; $N_I = N_T = N = 32$; ACI = -50 dB; Target SIR = 9 dB; Dynamic Range = 60 dB.

	Outage Probability	% Outage Due to Call Drops	% Outage Due to Call Blocks
without admission control	0.0104	44	56
with Algorithm I	0.0104	36	64
with Algorithm II	0.0107	23	77

system capacity did not improve with a larger dynamic range. We assume the system has an ACI specification of -50 dB to obtain a reasonable capacity. The data given in Table 2 shows that the outage probability under Algorithm II is lower compared to Algorithm I; hence Algorithm II is indeed a better algorithm. By using even more complicated algorithms, it may be possible to further reduce call dropping with a slight increase in outage probability.

3) Synchronization and System Simulation (J. Min, Prof. H. Samueli)

Robust synchronization algorithms for master-slave configured transceivers have been finalized. Acquisition in time, frequency, and hopping code is accomplished by means of frame synchronization. The control flow of this process is shown in Figure 1. It is achieved in two steps: coarse acquisition and fine acquisition. In our scheme, the receiver signal strength indicator (RSSI) output is used to coarsely indicate whether a C0 pilot tone is present or not (energy detection). The slave handset first listens for the C0 field of the frame, which is broadcast by the master at a pre-assigned acquisition carrier frequency. The RSSI output, which is generated by a cascade of logarithm amplifiers after the channel-

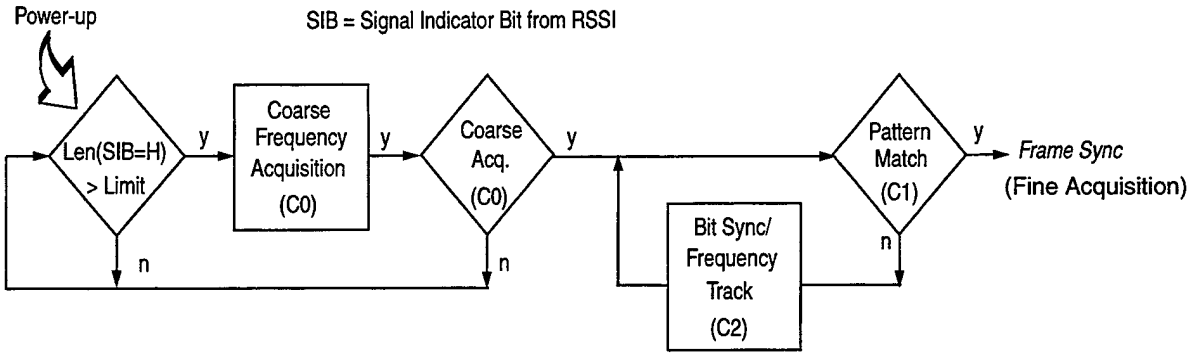


Figure 1. Frame synchronization control flow.

select filter, is then compared to a programmable energy threshold value to produce either high or low output. This signal indicator bit (SIB) lets the controller know whether a signal is present at a particular acquisition frequency. The energy threshold voltage is generated by an eight-bit digitally controlled sigma-delta DAC.

Once the bit goes high, frequency acquisition takes place by sweeping the frequency of the DDFS in a step of $F_{tone}/4$ (20 kHz). The frequency sweeping is required since the worst possible frequency error (±100 kHz) may be more than what the frequency tracking loop can handle. The five-bit correlation value of a pre-determined tone $(+F_{tone})$ is evaluated to indicate a coarse frequency lock. Simulation shows that a value of 17 out of the maximum possible 22 (for the case where the oversampling ratio N = 88) guarantees that the carrier frequency offset is within $\pm F_{tone}/4$ (20 kHz) from the desired. After coarse frequency acquisition, the rest of the C0 slots are used for coarse frame acquisition to detect a signal transiting from $+F_{tone}$ to $-F_{tone}$. The signal transition during the C0 slots still has $\pm T_{baud}$ uncertainty associated with it since bit synchronization has not been yet achieved. Therefore, the C1 (Receiver ID) pattern should be matched before frame synchronization is declared. Once a frame sync is declared (probably after the first frame), the receiver starts demodulating data. The pseudo-random (PN) code synchronization relies on the fact that once the frame is synchronized, a pre-defined frequency hopping pattern is repeated every frame. Thus, no extra PN code acquisition is required. If no match is detected over the window by the pattern recognizer, the synchronization circuit generates a *miss*. If more than five misses are reported, the receiver goes through a reacquisition process and start checking again the status of the SIB from the RSSI block.

The acquisition process only guarantees a phase accuracy within $T_{baud}/8$ (45 degrees) for the baud clock and an accuracy within $F_{tone}/4$ (20 kHz) for the carrier frequency. Thus, It is up to the tracking loop to pull in the remaining error and maintain its lock. For our receiver, both time tracking and frequency tracking loops use a second-order digital PLL architecture.

B. Architectures

1) RF/IF Architecture (A. Rofougaran, Profs. A. Abidi & H. Samueli)

An RF/IF architecture employing high-performance, low-power circuit techniques is developed for implementing the FH/SS transceiver. The combination of a direct digital frequency synthesizer (DDFS) and a digital-to-analog converter (DAC) is used to generate a single-sideband 26 MHz spread-spectrum waveform in the 902 to 928 MHz band. The signed I-Q frequency synthesis architecture reduces the highest frequency required from the DDFS/DAC to 13 MHz, yet covers the desired 26 MHz

hopping bandwidth. In addition, this single-sideband (SSB) modulation technique eliminates the need for image rejection filters with sharp roll-offs. Since we are using an FSK modulation scheme, a simple, highly efficient Class C differential power amplifier is used for transmission. Finally, a low-cost, off-the-shelf 902 to 928 MHz dielectric resonator bandpass filter is used between the power amplifier and the antenna to reject the out-of-band harmonics and meet FCC transmission mask requirements.

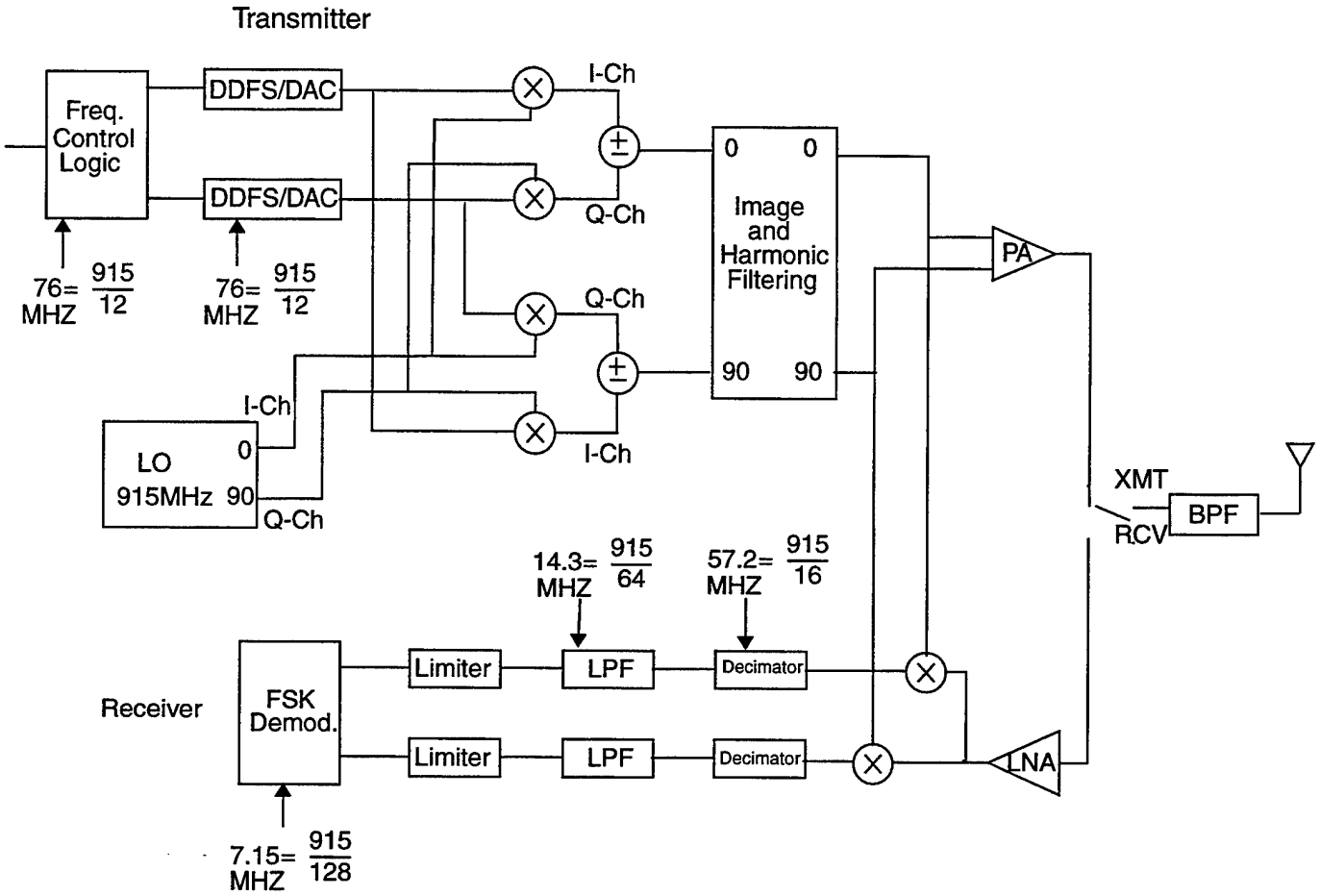


Figure 2. FH/SS Transceiver Architecture Block Diagram

Before down-conversion, the received signal must be amplified using a low noise amplifier (LNA). The gain and noise figure of the LNA are critical factors in determining the overall transceiver performance. The LNW with high gain at 915 MHz has been designed using monolithic high-Q inductors in CMOS without altering the standard fabrication process. The dehopping and down-conversion is performed in a single "direct-conversion" stage. Direct conversion requires accurate I-Q dehopping carriers at 915 MHz (Figure 2). These carriers are generated using the I-Q outputs of the DDFS/DAC and of the LO. They are then divided into lower frequencies to be used in the sub-sampling down-conversion mixers.

2) *Baseband Architecture (J. Min, Prof. H. Samueli)*

Low power dissipation requires that complex signal processing such as adaptive equalization and carrier recovery be avoided if possible. Delay spread without countermeasures (equalization) controls the upper limit of the transmission rate through the radio channel. For outdoor microcellular channels, the value is around 0.5 μs. Our modulation baud rate (80 kHz) is appropriately chosen for this requirement.

Thus, no equalization is required for our receiver. Noncoherent detection also avoids complex carrier recovery. Automatic gain control (AGC) is required if the dynamic range of the detector is less than that encountered in the radio channel, typically 80 dB. Since FSK signalling only requires the frequency information, not the amplitude, a hard-limiter capable of more than 80 dB of dynamic range replaces an actual AGC loop. Due to odd harmonics produced by hard-limiting the sine waves, our approach limits the modulation scheme to binary FSK for the real signal and to quaternary FSK for the complex signal. However, the merits from power savings and hardware simplicity obtained from the absence of a multi-bit analog-to-digital converter (ADC) and an expensive (in both power and complexity) linear variable gain amplifier required in the AGC loop justify this architecture for a portable transceiver design. This receiver architecture results in an all-digital quadrature FSK demodulator using an over-sampled one-bit correlation detection scheme.

The remaining baseband processing blocks preceding the modulator or following the demodulator include an interleaver/deinterleaver, a hop repeater/combiner, a channel coder/decoder, an interface for voice and data, and a controller. The baseband processing architecture for a prototype handset, which supports both voice and serial data, has been developed in this research project (Figure 3).

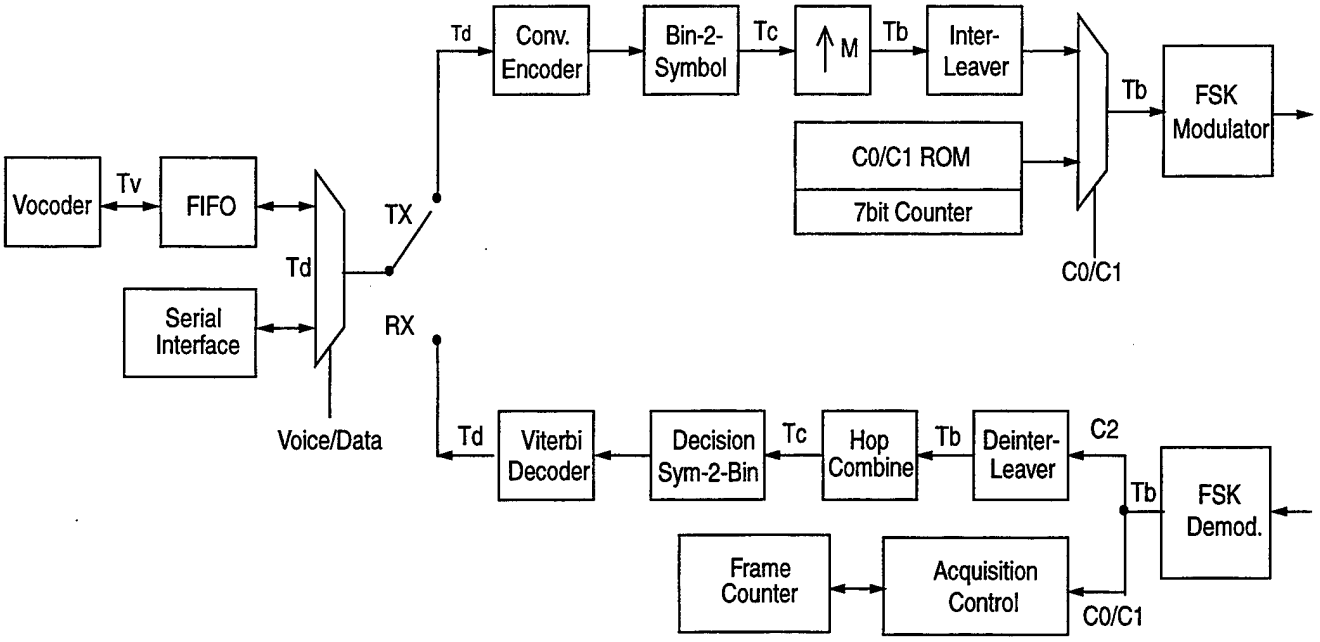


Figure 3. Baseband Processing Block Diagram

C. Advanced Analog and Digital Circuits

1) Spread-Spectrum RF to Baseband Receiver (J. Chang, J.Y. Chang, S. Khorram, M. Djafari, and Prof. A. Abidi)

The receiver consists of a low-noise-amplifier (LNA), parallel I and Q channels of mixer, lowpass filter, and limiter. The LNA utilizes an on-chip suspended spiral inductor for tuning and noise shaping. It is a wideband common-gate stage which also provides input mathing to 50 Ω The LNA provides 20 dB of gain while drawing 4.4 mA of current. The double-side band (DSB) noise figure of the LNA is 3.5 dB. The LNA and mixer combination is highly linear, providing a combined gain of 34 dB with -6 dBm of

input third-order intermodulation intercept point (IIP3). The LNA-mixer combined DSB noise figure is 5.5 dB.

The low-pass filter is a switched-capacitor, sixth-order elliptic filter. The filter is implemented using 3 bi-quad sections which are optimized for high dynamic range and low-noise performance. It dissipates 20 mW from a 3V supply with 24 dB of voltage gain. The limiter is a cascade of seven identical amplifiers, each with a gain of 4. This saturates the incoming signal and produces a square-wave at the output. By summing equally weighted taps at the outputs of the seven amplifiers, the logarithmic function is synthesized as a piecewise linear function. This logarithmic function output serves as the received signal strength indicator (RSSI).

The receiver has been fabricated and tested with the transmitter on one chip using 1-μm CMOS. It has an ensemble noise figure of 8.5 dB, which enables the receiver to detect -105 dBm signals at 10^{-6} bit-error-rate. Its input IP3 is -8.5 dBm, and thus, its dynamic range is 96.5 dB. The receiver draws 120 mA with a single 3-volt supply.

2) *Local Oscillator (PLL/VCO) (J. Rael, Prof. A. Abidi)*

A low phase noise clock source is required to up-convert and down-convert the modulated data in the transceiver. However, the required clock frequency is much higher than the frequencies typically attainable with conventional crystal oscillators. Therefore, a frequency synthesizer (FS) is needed which can generate these high frequencies by scaling a stable reference source. The FS must not only have low phase noise, but low power dissipation as well. These specifications are further complicated by wide temperature range and low-power supply voltage. A phase-locked loop (PLL) was chosen

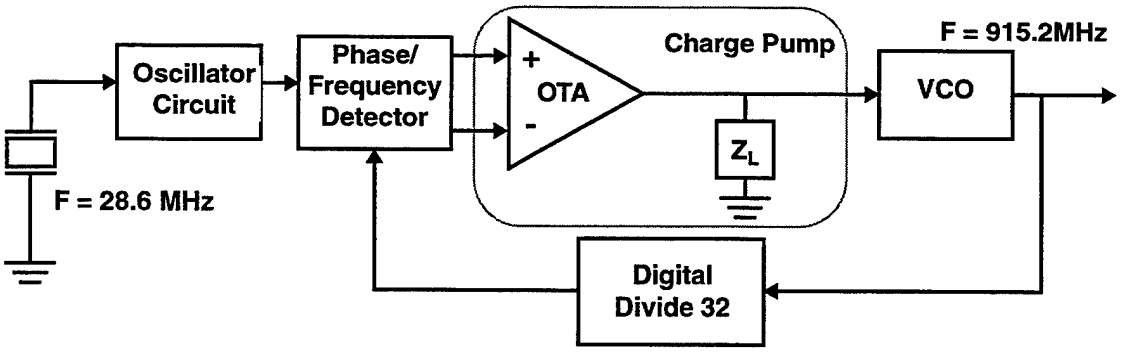


Figure 4. PLL Block Diagram

because it can be easily integrated, and provides low power dissipation while keeping frequency spurs to a minimum.The block diagram of the frequency synthesizer is given in Figure 4.

The optimized four-stage ring oscillator was tested and the phase-noise was measured to be -90 dB/Hz at 100 kHz offset. This oscillator consumed 70 mW of power but oscillated at only 660 MHz.

To improve the phase noise and increase the oscillating frequency, we have developed a new oscillator topology. This new topology consists of two oscillators, coupled such that quadrature outputs are generated. The first oscillator using this technique employed etched inductors as load devices and had a measured phase noise of -101 dB/Hz at 100 kHz offset with a power dissipation of only 40 mW. The error in quadrature was estimated to be less than 1° by measuring the rejection of the unwanted sideband in a single sideband modulator. The center frequency of the oscillator was 928 MHz.

The second oscillator used simulated inductors as load devices. This oscillator occupied much less area but had higher phase noise than the etched inductors. The measured phase noise was -92 dB/Hz

at 100 kHz offset with a power dissipation of 90 mW. The error in quadrature was estimated to be less than 1° with a center frequency of 866 MHz.

3) Power Amplifier (M. & A. Rofougaran, Prof. A. Abidi)

A MHz RF power amplifier with a digital power control has been designed and implemented in standard 1-μm CMOS process to be used in a frequency-hopped spread-spectrum transceiver. The power amplifier consists of a low biased switching mode output stage with an inductive load preamplifier that employs a 75 nH on-chip suspended inductor. The large on-chip inductor exploits a large gain at the resonant frequency with low power consumption. The fabricated IC chip has been measured, characterized and seems to closely match the simulated results and design specifications.

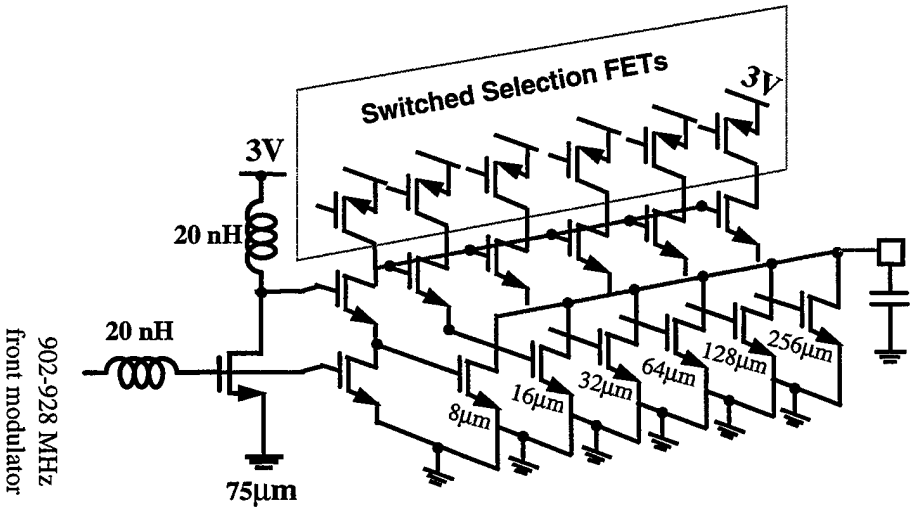


Figure 5. Modified Version of the PA with Two Unetched 20 nH Inductors

The power amplifier exploits a 40% percent efficiency The power amplifier is capable of transmitting 6uW to 20mW using digital power control. This corresponds to a 35 dB dynamic range. The circuit shows low linearity because of low biasing input and hard drive of the input.

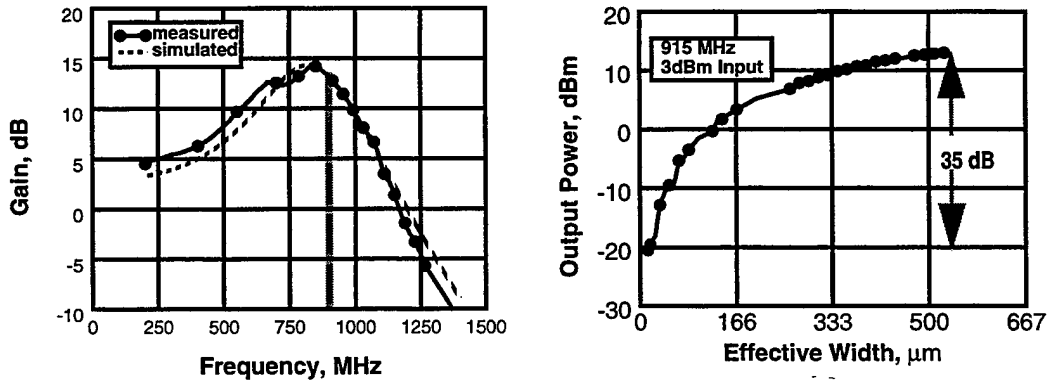


Figure 6. Low biasing input and hard drive output

4) *Spread Spectrum Baseband to RF Transmitter (G. Chang, Prof. A. Abidi)*

A monolithic spread spectrum baseband-to-RF transmitter has been designed in 1-μm standard digital CMOS process. This transmitter (see figure below) consists of two agile baseband I-Q frequency synthesizers, a low-noise 915 MHz local oscillator (LO), two upconversion mixers, an on-chip RF LC bandpass filter, and a power amplifier. In the second design iteration, a passive RC polyphase filter has been implemented, replacing the need for an LC bandpass filter which occupies a large silicon area. The transmitter uses a signed I-Q architecture with single-sideband (SSB) modulation technique that eases the design specifications of the frequency synthesizer. The baseband data signals are up-converted to RF frequencies in a single-shot conversion, thus eliminating the need for any IF circuits. By employing such a direct-conversion architecture and adequate frequency planning, this single-chip transmitter incorporates all the functions from baseband data into an antenna drive without any off-chip components such as discrete inductors or SAW filters.

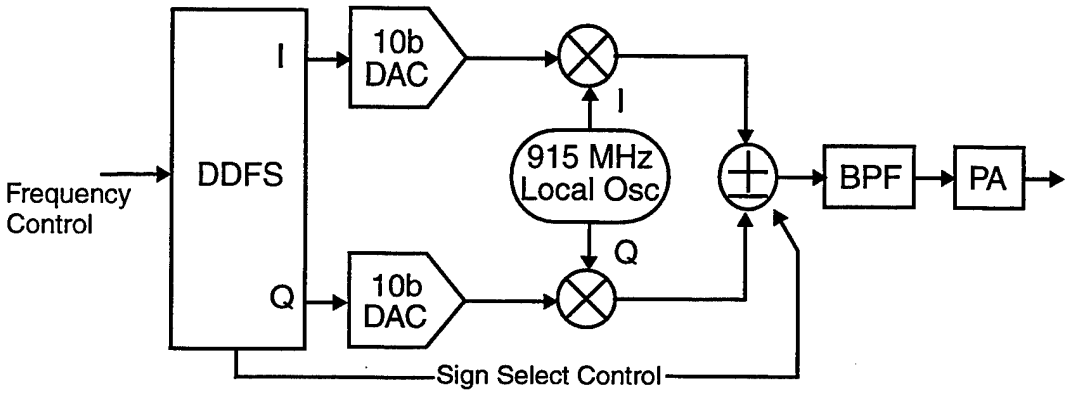


Figure 7. RF Transmitter

The frequency synthesizer is implemented using a direct digital frequency synthesizer (DDFS) with a 24-bit frequency-control word followed by a low-power 10-bit digital-to-analog converter (DAC) employing pipelined switched-capacitor charge-redistribution algorithm. The LO contains on-chip inductor loads, and its architecture inherently generates both quadrature and differential outputs. Each upconversion mixer is a passive implementation, comprising a four-FET commuting switch. The on-chip LC filter is made by using an etched inductor in series with the on-chip parasitic capacitance. The power amplifier is capable of delivering variable output power into a 50 Ω antenna ranging from 20 μW to 20 mW under the control of a 6-bit digital control word.

Conventionally, a continuous-time low-pass filter is required to suppress these images to prevent them from propagating further down the system and creating intermodulations in the band of interest. In our implementation of the transmitter IC, however, the need for this low-pass filter is eliminated as a result of clever frequency planning of the distribution of DAC image frequencies. In this fashion, with the appropriate allocation of unavoidable image frequencies and inter-modulated tones, the task of a continuous-time low-pass filter is shifted to the off-chip resonator filter which is needed to filter out-of-band images regardless of the system architecture.

The fabricated transmitter has been fully characterized. The power amplifier delivers measured output power ranging from -17 dBm (20 μW) to +14 dBm (25 mW). At the maximum rated power, the measured transmitter output spectrum shows the unwanted sideband at -42 dBc, and the LO feedthrough at -48 dBc. The achieved-42 dBc image sideband suppression implies that the overall on-chip quadrature inaccuracy is less than 1°. The transmitter portion drains 100 mA in active mode from a 3 V power supply, except for an RF buffer in the power amplifier section, which requires 3.5 V.

5) *Digital Baseband Receiver (H.-C. Liu, Prof. H. Samueli)*

A baseband receiver IC has been designed and fabricated using the HP 1-μm CMOS process (Figure 8). The chip performs noncoherent binary/quaternary FSK demodulation, equal-gain diversity combining of dual antenna branches, and symbol and frequency synchronization. Since the target application is for a handheld transceiver, both low power and small area are important considerations. Several techniques are applied in different levels of the design to achieve this goal. Since the input is only one bit, the local tone generator is also implemented using a numerically controlled oscillator (NCO), rather than a sophisticated direct digital frequency synthesizer (DDFS). The demodulator supports both binary and quaternary FSK signalling. The binary format consists of tones from $\{+F_{tone}, -F_{tone}\}$, while the quaternary format contains tones from $\{-2F_{tone}, -F_{tone}, +F_{tone}, +2F_{tone}\}$. Thus, complex $\pm F_{tone}$ and $\pm 2F_{tone}$ detectors are required. In addition, a $\pm F_{tone}/2$ correlator is used in the frequency tracking loop. The quadrature NCO generates a 1-bit I-Q $2F_{tone}$ square-wave; the other reference tones, F_{tone} and $F_{tone}/2$, are easily produced by simple binary division of the NCO tone. A noncoherent (NC) demodulator requires a magnitude calculation unit, which is conventionally implemented with a pair of squaring multipliers followed by a summing node. In our architecture, an absolute value addition block replaces this squaring block. Thus, a truly multiplierless NC FSK demodulator has been implemented with little performance degradation. Symbol timing must be derived from the received signal in order to synchronously sample the output of correlators at the proper time instant. An early-late gate synchronizer has been designed to control the clock timing of the integrate-and-dump (I/D) units in the tone detector. The PN code generator provides the frequency hopping pattern which is fine-tuned by the output of the loop filter, and controls the frequency word of the DDFS/DAC block. The hop time synchronization is performed in an open-loop fashion, aligning the hop time with the baud time plus a programmable delay to compensate for the delay between the two loops. Parallelism is employed in the blocks with a high clock rate to eliminate the power consumption of time multiplexing and control. On the other hand, time-multiplexed implementation is applied on the blocks with a low clock rate to minimize the chip area. A robust, gated clocking scheme is used for registers with a low access rate to reduce the power consumption.The fabricated baseband receiver chip was found to be fully functional with a power dissipation of 4.5 mW from a 3-V supply. The rest of baseband signal processing including frame acquisition and data interface was developed in Xilinx FPGA chips.

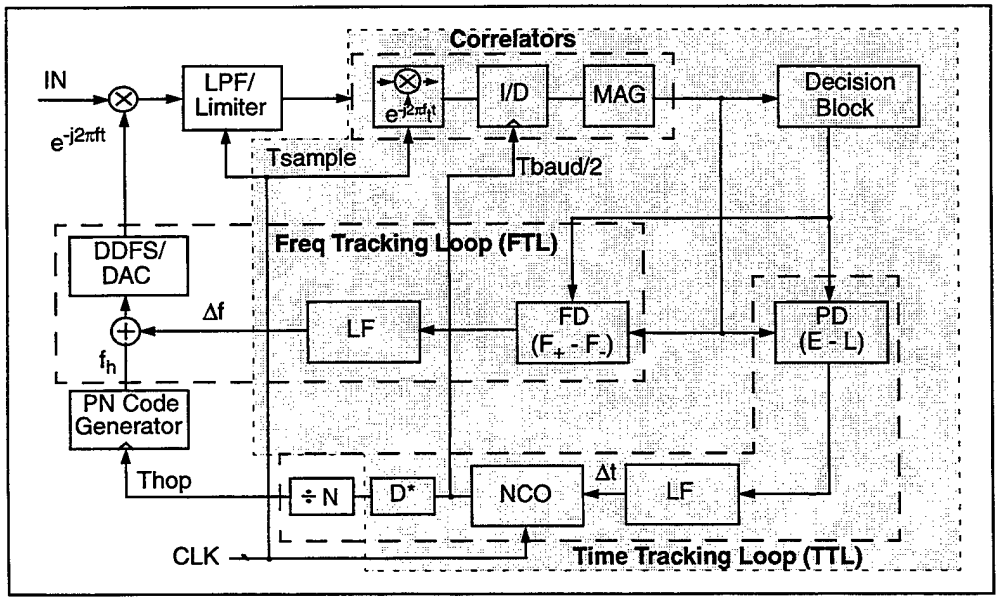


Figure 8. Baseband receiver chip block diagram

D. Miniature Antenna Design

1) Analysis and Design Techniques (J. Colburn, Prof. Y. Rahmat-Samii)

The intent of the experimental work involved in miniature antenna design was to evaluate the diversity performance of three different dual antenna configurations on a handset chassis. Measurements have been made for characterizing the fading band-width for 902-928 indoor channel.

Figure 9 illustrates the three different dual antenna configurations proposed for this project: (a) two planar inverted-F antennas (PIFA) mounted on the sides of a handset chassis; (b) a PIFA mounted on the top of the handset chassis and a monopole integrated into the "flip" of the handset; and (c) a PIFA mounted on the back and a helix mounted on the top of a handset chassis. All antennas, except the helix, were simulated on the handset chassis using a finite-difference time-domain (FDTD) computer program, and were designed to achieve better than a 2:1 voltage standing wave ratio (VSWR) to 50 ohm line over the 902-928 MHZ bandwidth in addition to good diversity performance. Various indoor propagation experiments were completed (under FCC experimental license 4660-EX-PL-94) for no line-of-sight paths in this evaluation. Figure 10 is a plot of a short segment of data recorded with the two side mounted PIFA design while transversing an indoor test path. Plotted is the signal power received simultaneously over the two antennas. Averages of the observed power correlation coefficients are: 0.5 for the two side mounted PIFA configuration; 0.6 for the top mounted PIFA/filp monopole configuration; 0.6 for the back mounted PIFA/helix configuration. These measured results indicate good diversity performance for all three proposed designs.

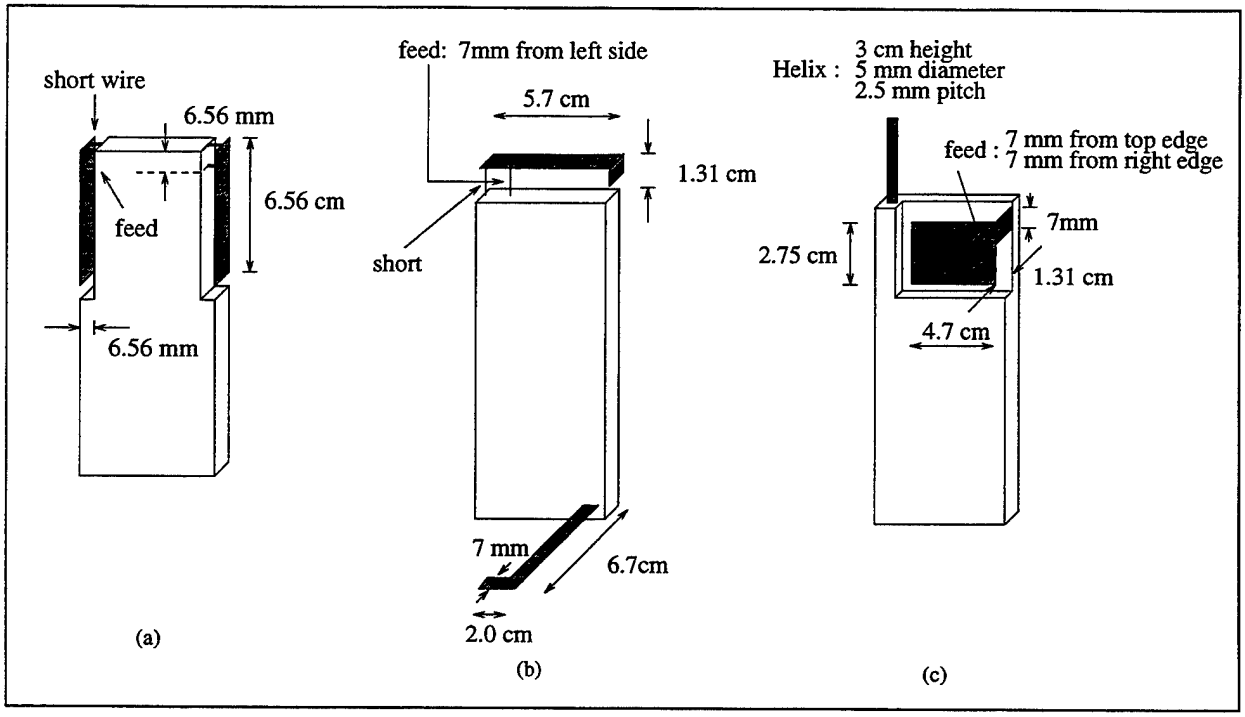


Figure 9. Schematics of handset: (a) side mounted PIFAs; (b) top mounted PIFA/"flip" monopole; (c) back mounted PIFA/helix

Experiments to characterize the fading bandwidth of the 902-928 MHZ indoor channel were also completed. Correlation bandwidths between 2 to 8 MHz were observed.

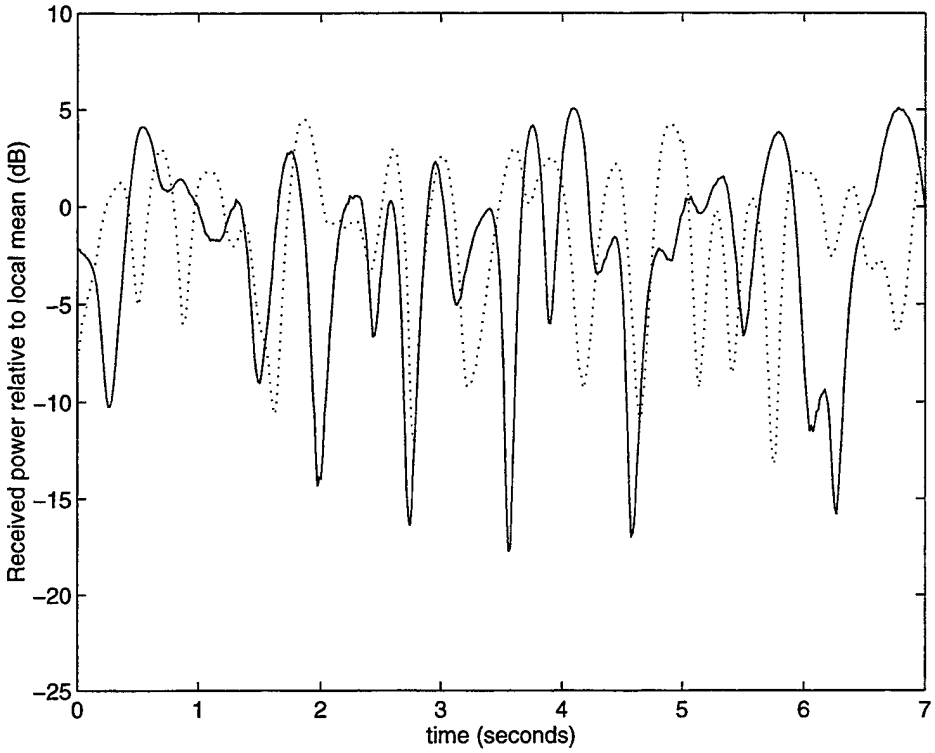


Figure 10. Sample of signals recorded with two side mounted PIFA antenna configuration

III. Presentations and Publications

[1] J.S. Colburn, Y. Rahmat-Samii, M.A. Jensen, and G.J. Pottie, "Diversity Performance of Dual Antenna Personal Communication Handsets", *Proceedings of the IEEE Antennas and Propagation Society International Symposium,* Baltimore, MD, vol. 1, pp. 730-733, July 1996. **[p. V.1 of this report]**

[2] P.J. Chang, A. Rofougaran, and A.A. Abidi, "A CMOS Channel-Select Filter for a Direct-Conversion Wireless Receiver", *Proceedings of the IEEE Symposium on VLSI Circuits*, Honolulu HI, pp. 62-63, June 1996; submitted for publication to the *IEEE Journal of Solid-State Circuits.* **[p. V.2 of this report]**

[3] R.A. Hawley, B.C. Wong, T. Lin, J. Laskowski, and H. Samueli, "Design Techniques for Silicon Compiler Implementations of High-Speed FIR Digital Filters", *IEEE Journal of Solid-State Circuits*, vol. 31, no. 5, pp. 656-667, May 1996. **[p. V.3 of this report]**

[4] A. Rofougaran, G. Chang, J.J. Rael, M. Rofougaran, S. Khorram, M.-K. Ku, E. Roth, A.A. Abidi, and H. Samueli, "A 900 MHz CMOS Frequency-Hopped Spread-Spectrum RF Transmitter IC", *Proceedings of the IEEE Custom Integrated Circuits Conference*, San Diego CA, pp. 209-212, May 1996. **[p. V.4 of this report]**

[5] A. A. Abidi, A. Rofougaran, G. Chang, J. Rael, J. Chang, M. Rofougaran, P.J. Chang, and S. Khorram, "A Monolithic 900 MHz Spread-Spectrum Wireless Transceiver in 1-μm CMOS", presented at the Workshop on Advances in Analog Circuit Design (AACD), Lausanne Switzerland, April 1996 (proceedings limited to workshop participants). **[p. V.5 of this report]**

[6] H.-C. Liu, J. Min and H. Samueli, "A Low-Power Baseband Receiver IC for Frequency-Hopped Spread Spectrum Applications", *Proceedings of the IEEE Custom Integrated Circuits Conference*, Santa Clara CA, pp. 311-314, May 1995; *IEEE Journal of Solid-State Circuits,* vol. 31, no. 3, pp. 384-394, March 1996. **[p. V.6 of this report]**

[7] L.K. Tan, E. Roth, G. Yee, and H. Samueli, "An 800-MHz Quadrature Digital Synthesizer with ECL-Compatible Output Drivers in 0.8 μm CMOS", *IEEE Journal of Solid-State Circuits*, vol. 30, no. 12, pp. 1463-1473, December 1995. **[p. V.7 of this report]**

[8] A. Rofougaran, J. Y-C. Chang, M. Rofougaran, S. Khorram, and A. A. Abidi, "A 1 GHz CMOS RF Front-End IC with Wide Dynamic Range", *Twenty-first European Solid-State Circuits Conference Proceedings*, pp. 250-253, September 1995. **[p. V.8 of this report]**

[9] J.S. Colburn and Y. Rahmat-Samii, "Analysis of a Square Helix Applicable to the Personal Satellite Communication Handset", *Proceedings of the IEEE Antennas and Propagation Society International Symposium*, Newport CA, vol. 3, pp. 1418-1421, June 1995. **[p. V.9 of this report]**

[10] S. Khorram, A. Rofougaran, and A.A. Abidi, "A CMOS Limiting Amplifier and Signal-Strength Indicator", *1995 Symposium on VLSI Circuits Digest of Technical Papers,* pp. 95-96, June 1995. **[p. V.10 of this report]**

[11] J. Min, H.-C. Liu, A. Rofougaran, S. Khorram, H. Samueli, and A.A. Abidi, "A Low-Power Correlation Detector for Binary FSK Direct-Conversion Receivers", *Electronics Letters,* vol. 31, no. 13, pp. 1030-1032, June 1995. **[p. V.11 of this report]**

[12] J. Min, "Analysis and Design of a Frequency-Hopped Spread Spectrum Transceiver for Wireless Personal Communications", UCLA Ph. D. dissertation, June 1995.

[13] J. Rael, "A 915 MHz CMOS Frequency Synthesizer", UCLA Masters Thesis, June 1995.

[14] M. Rofougaran, "A MHz RF Power Amplifier in 1-μm CMOS for a Spread-Spectrum Communication Transceiver", UCLA Masters Thesis, June 1995.

[15] Abidi, A.A., "Low Power Radio-Frequency ICs for Portable Communications", *Proceedings of the IEEE*, vol. 83, no. 4, pp. 544-569, April 1995. (Invited Paper) **[p. V.12 of this report]**

[16] M.A. Jensen and Y. Rahmat-Samii, "EM Interaction of Handset Antennas and a Human in Personal Communications", *Proceedings of the IEEE*, vol. 83, no. 1, pp. 7-17, Jan. 1995. **[p. V.13 of this report]**

[17] G.J. Pottie and A.R. Calderbank, "Channel Coding Strategies for Cellular Radio", *IEEE Transactions on Vehicular Technology*, vol. 44, no. 4, pp. 763-770, 1995 (also presented at the IEEE International Symposium on Information Theory, San Antonio TX, Jan. 1993) **[p. V.14 of this report]**

[18] M.A. Jensen and Y. Rahmat-Samii, "Performance Analysis of Antennas for Hand-Held Transceivers Using FDTD", *IEEE Transactions on Antennas and Propagation*, vol. 42, no. 8, pp. 1106-1113, August 1994. **[p. V.15 of this report]**

[19] A. Abidi, "Low Power RF-ICs in Wireless Transceivers", *Proceedings 1994 IEEE Symposium on Low Power Electronics*, pp. 18-21, August 1994. **[p. V.16 of this report]**

[20] J.S. Colburn, M.A. Jensen, and Y. Rahmat-Samii, "Comparison of MOM and FDTD for Radiation and Scattering Involving Dielectric Objects", *Proceedings of the IEEE Antennas and Propagation Society International Symposium*, Seattle WA, pp. 1802-1805, June 1994. **[p. V.17 of this report]**

[21] M.A. Jensen and Y. Rahmat-Samii, "The Electromagnetic Interaction Between Biological Tissue and Antennas on a Transceiver", *Proceedings of the IEEE Antennas and Propagation Society International Symposium*, pp. 367-370, Seattle WA, June 1994. **[p. V.18 of this report]**

[22] M. Rofougaran, A. Rofougaran, C. Olgaard, and A.A. Abidi, "A 900 MHz CMOS RF Power Amplifier with Programmable Output", *1994 Symposium on VLSI Circuits Digest of Technical Papers*, pp. 133-134, June 1994. **[p. V.19 of this report]**

[23] A. Abidi, "Radio-Frequency Integrated Circuits for Portable Communications", *Proceedings of the IEEE 1994 Custom Integrated Circuits Conference*, San Diego CA, pp. 151-158, May 1994. **[p. V.20 of this report]**

[24] J. Min, A. Rofougaran, H. Samueli, and A. Abidi, "An All-CMOS Architecture for a Low-Power Frequency-Hopped 900 MHz Spread Spectrum Transceiver", *Proceedings of the IEEE 1994 Custom Integrated Circuits Conference*, San Diego CA, pp. 379-382, May 1994. **[p. V.21 of this report]**

[25] M.A. Jensen and Y. Rahmat-Samii, "Characterisation of Electromagnetically Coupled Superquadric Loop Antennas for Mobile Communications Applications", *IEEE Proceedings - Microwaves, Antennas and Propagation*, vol. 141, no. 2, pp. 85-93, April 1994. **[p. V.22 of this report]**

[26] M.A. Jensen and Y. Rahmat-Samii, "Characterisation of Antennas for Personal Wireless Communications Applications", *International Journal of Wireless Information Networks*, vol. 1, no. 3, March 1994. **[p. V.23 of this report]**

[27] G. Chang, A. Rofougaran, M.-K. Ku, A.A. Abidi, and H. Samueli, "A Low-Power CMOS Digitally Synthesized 0-13 MHz Sinewave Generator", *IEEE 1994 International Solid-State Circuits Conference Digest of Technical Papers*, pp. 32-33 (slide supplement pp. 20-21), Feb. 1994. **[p. V.24 of this report]**

[28] V.S. Lin and G.J. Pottie, "Channel Coding for a Frequency-Hopped Wireless Transceiver", *Lecture Notes in Computer Science*, A.T. Goldberg and N.P. Secord, eds., pp. 109-124, Springer-Verlag, 1994 (presented at the Third Canadian Workshop on Information Theory, Rockland Ontario, May 1993). **[p. V.25 of this report]**

[29] J. Min, A. Rofougaran, V. Lin, M. Jensen, H. Samueli, A. Abidi, G. Pottie, and Y. Rahmat-Samii, "A Low-Power Handheld Frequency-Hopped Spread Spectrum Transceiver Hardware Architecture",

Wireless Personal Communications, Trends, and Challenges, chapter 11, pp. 119-126, Kluwer Academic Publishers, 1994 (presented at the IEEE Workshop on VLSI in Communications, Lake Tahoe CA, Sept. 1993) **[p. V.26 of this report]**

[30] P.Y. Chan, A. Rofougaran, K.A. Ahmed, and A.A. Abidi, "A Highly Linear 1-GHz CMOS Downconversion Mixer", *Nineteenth European Solid-State Circuits Conference Proceedings,* Seville Spain, pp. 210-213, Sept. 1993. **[p. V.27 of this report]**

[31] M.A. Jensen and Y. Rahmat-Samii, "Isolated and Coupled Superquadric Loop Antennas for Mobile Communications Applications", *Proceedings of the Third International Mobile Satellite Conference*, Pasadena CA, pp. 581-586, June 1993. **[p. V.28 of this report]**

[32] M.A. Jensen and Y. Rahmat-Samii, "Characteristics of Superquadric Loop Antennas for Mobile Communications Applications", presented at the IEEE Antennas and Propagation Society International Symposium, Ann Arbor MI, June 1993. **[p. V.29 of this report]**

[33] M.A. Jensen and Y. Rahmat-Samii, "FDTD Analysis of PIFA Diversity Antennas on a Handheld Transceiver Unit", presented at the IEEE Antennas and Propagation Society International Symposium, Ann Arbor MI, June 1993. **[p. V.30 of this report]**

[34] J. Y.-C. Chang, A.A. Abidi, and M. Gaitan, "Large Suspended Inductors on Silicon and Their Use in a 2-μm CMOS RF Amplifier", *IEEE Electron Device Letters*, vol. 14, no. 5, pp. 246-248, May 1993. **[p. V.31 of this report]**

[35] J.Y.-C. Chang and A.A. Abidi, "A 750 MHz RF Amplifier in 2-μm CMOS", *1993 Symposium on VLSI Circuits Digest of Technical Papers*, Kyoto Japan, pp. 111-112, May 1993. **[p. V.32 of this report]**

[36] M.A. Jensen and Y. Rahmat-Samii, "Superquadric Loop Antennas", presented at the IEEE International Conference on Antennas and Propagation, Edinburgh Scotland, UK, March 1993. **[p. V.33 of this report]**

[37] M.A. Jensen and Y. Rahmat-Samii, "Analysis of Isolated and Coupled Superquadric Loop Antennas", presented at the URSI National Radio Science Meeting, Boulder CO., Jan. 1993 (abstract published in the United States Natinal Committee International Union of Radio Science *Program and Abstracts: National Radio Science Meeting*, p. 9, Jan. 1993). **[p. V.34 of this report]**

[38] J.Y.-C. Chang and A.A. Abidi, "Fabrication Technique for Large Inductors Usable at High Frequencies on Semiconductor Substrates", **Patent Application** filed by the University of California, Los Angeles, 1995.

IV. Personnel

Faculty:

Henry Samueli, Professor	Principal Investigator
Asad A. Abidi, Professor	Co-Principal Investigator
Gregory J. Pottie, Associate Professor	Co-Principal Investigator
Yahya Rahmat-Samii, Professor	Co-Principal Investigator

Staff:

Irene Andreadis, Administrative Analyst

Mandy Tanner, Administrative Analyst

Edward W. Roth, ICS Laboratory Manager

Graduate Student Researchers:

Victor Lin	Joseph Colburn	Jonathan Min
Jacob Rael	Mong-Kai Ku	Glenn Chang
Ahmadreza Rofougaran	Maryam Rofougaran	James Y.C. Chang
Huan-Chang Liu	Michael Jensen	

Technical Point of Contact:

Prof. Henry Samueli
Electrical Engineering Department
University of California, Los Angeles
56-125B Engineering IV Box 951594
Los Angeles, CA 95-1594
Tel: (310) 825-4854
FAX: (310) 825-7928
E-mail: samueli@janet.ucla.edu

Administrative Contact

Ms. Mandy Tanner
Electrical Engineering Department
University of California, Los Angeles
56-125B Engineering IV Box 951594
Los Angeles, CA 95-1594
Tel: (310) 206-7961
FAX: (310) 825-7928
E-mail: mandy@janet.ucla.edu

V. Appendix: Reprints of Publications

Diversity Performance of Dual Antenna Personal Communication Handsets

J.S. Colburn, Y. Rahmat-Samii, M.A. Jensen, and G.J. Pottie

Proceedings of the IEEE Antennas and Propagation Society International Symposium
vol. 1, pp. 730-733
June 1996

Diversity Performance of Dual Antenna Personal Communication Handsets

J. S. Colburn*[,1,2], Y. Rahmat-Samii[1], M. A. Jensen[3] and G. J. Pottie[1]

[1]Department of Electrical Engineering
University of California, Los Angeles
Los Angeles, CA, 90095-1594

[2]TRW Space and Electronics Group
One Space Park
Redondo Beach, CA, 90278

[3]Department of Electrical Engineering
Brigham Young University
Provo, UT, 84602

I. INTRODUCTION Short-term fading due to multipath propagation has a significant impact on personal wireless communication systems. Mitigating signal strength fluctuations due to interference between multiple replicas of the signal of interest arriving at the receiver over different paths is an active area of research. Antenna diversity is the technique where multiple antennas are used for transmission/reception to help overcome the effects of multipath fading [1,2]. This paper gives a brief description of experiments, data processing and results used to evaluate the diversity performance of three candidate dual antenna handset configurations: two side mounted planar inverted F antennas (PIFA); back mounted PIFA with top mounted helix; top mounted PIFA and "flip" monopole [3]. The indoor ISM band (902 - 928 MHz) propagation channel was of interest. The issue of data normalization is addressed and results showing its effect on observed correlation are given. It is seen that the three proposed diversity systems yield sufficient decorrelation to warrant consideration for use in diversity systems.

II. ANTENNA CONFIGURATIONS This work was undertaken to evaluate the diversity performance of three dual antenna configurations for a small transceiver handset. Figure 1 illustrates the configurations considered: (a) top mounted helix and back mounted planar inverted F antenna (PIFA); (b) side mounted PIFAs; (c) top mounted PIFA and flip monopole. The overall dimensions of the handset geometry considered are given in Figure 1 (c), with the specific antenna locations and dimensions for each design specified in Figures 1 (a), (b) and (c).

The dual antenna handsets illustrated in Figure 1 achieve their overall diversity performance from all three types of antenna diversity: spatial, polarization and pattern. Consider the top mounted helix and back mounted PIFA structure shown in Figure 1 (a). These two antennas are displaced from each other (approximately 5 cm) and have different pattern and polarization characteristics, resulting in significant decorrelation between the two received signals. Similarly, the dual side mounted PIFA configuration illustrated in Figure 1 (b) achieves pattern and polarization diversity through physical element displacement as well as opposite orientation [4]. These same descriptions also apply to the top mounted PIFA and "flip" monopole handset of Figure 1 (c), although in this case a significantly larger element separation is used (approximately 15 cm).

III. EXPERIMENTAL SETUP and PROCEDURE The transmitter for these tests consisted of a sweep oscillator connected to a $\lambda/2$ dipole antenna. The

source was operated in single-frequency mode at 915 MHz. On the receiver end, the outputs of the dual antennas under test were amplified, then piped into separate spectrum analyzers. The spectrum analyzers were put in zero span mode and centered on 915 MHz, for which the video output signals were proportional to the received power in dBm. The analog video output signals of the spectrum analyzers were then digitized and the data recorded on a personal computer. The digitizing rate used was approximately 72 samples per second per channel.

For each test run, the receiver was located in a typical laboratory room in the Engineering IV building on the UCLA campus while the transmitting antenna was moved along one of six predetermined paths at a constant speed. The paths were chosen such that there was no line-of-sight path. Paths 1, 2 and 6 were located on the same floor as the receiver, while paths 3, 4 and 5 were located on the floor below the receiver. The Engineering IV building on the UCLA campus is a large office building (approximate floor dimensions are 67 m x 60 m) of modern construction, consisting of individual offices, large open laboratories and cubical areas. During each test run the transmitter was moved approximately 0.5 m/s resulting in approximately 47 samples per wavelength. The time length of each test run was 30 seconds.

IV. RESULTS In addition to the dual antenna geometries described prior, the diversity performance of two vertically orientated $\lambda/2$ dipoles with horizontal separation of 0.06λ and 0.4λ (at 915 MHz) were measured for reference purposes. Figures 2 (a) and (b) are plots of different segment lengths of raw data recorded on one of the test paths from one of the two dipole antennas which were spaced $0.06\ \lambda$ apart. In Figure 2 (a) both the long-term shadow/path-loss fading and short-term multipath fading can be seen. In Figure 2 (b), which has an expanded time scale, the short-term multipath fading and a very fast ripple that is believed to be experimental noise can be seen. To determine the true diversity performance of the dual antenna handsets, both the long-term shadow/path-loss fading and experimental noise had to be removed from the measured data, leaving just the multipath fading behavior. In this work the long-term shadow/pathloss fading was assumed to be a multiplication factor of the observed envelope and was extracted by normalizing the raw data with the *local* mean. Various time lengths were used in the determination of the *local* mean and their effect on the calculated correlation coefficient recorded. The experimental noise was removed with a very short term averaging. The number of points in this averaging was varied and effects recorded. Figures 2 (c) and (d) are plots of the same data segments shown in Figures 2 (a) and (b) after a three point smoothing average was used to remove the very fast ripple and the local mean (calculated using a 2 second moving average) was extracted.

The cumulative distribution functions (CDFs) of the measured data sets were seen to fit the Rician statistic well [1,3]. Curve fitting between the measured CDFs and theoretical Rician distributions of various dominate ray to multipath signal power ratios was used to determine the multipath severity of each path considered. Based on these results, data from test paths exhibiting a strong dominant ray path were eliminated from further consideration since it was intended to concentrate on the worst case multipath environment.

Table 1 lists the power correlation coefficients computed from the data set for the

dual side mounted PIFA handset recorded over one of the test paths. In the table the time period used in the calculation of the *local* mean and the number of points used in the short term averaging are parameters that are varied. Similar results were seen in all the data sets [3]. With the exception of very short *local* mean periods, the computed values of the correlation coefficients were seen to be fairly insensitive to the number of points used in the short term averaging for removal of the experimental noise and the time period used to calculate the *local* mean for removal of the shadow/path-loss fading. (Very short time periods in the *local* mean calculation result in normalizing out the fast fading behavior, thus yielding higher correlation coefficients.)

Table 2 lists the computed correlation coefficients for the five different dual antenna configurations noted earlier for the test paths considered most relevant. The data shown in Table 2 assumes a 2 second demeaning and 3 point short-term averaging of the raw data. From Table 2 one can see all three proposed dual antenna handset designs achieve diversity performance equal to or greater than that of two vertically orientated dipole antennas with 0.4 λ of horizontal separation. With maximal-ratio combining in a Rayleigh fading environment, it is possible to achieve 10 of the possible 12 dB of diversity gain of a dual branch system over a single channel system (assuming a 99% signal reliability criterion) with 0.6 branch correlation [2]. This fact and the data from these experiments indicate that the three proposed dual antenna handset configurations achieve sufficient decorrelation to warrant their consideration for use in a diversity system.

Acknowledgments: This work was completed at UCLA and supported by ARPA Contract DAAB07-93-C-C501.

REFERENCES

[1] D. Parsons, *The Mobile Radio Propagation Channel*, 1992.
[2] W. C. Lee, *Mobile Communications Engineering*, 1982.
[3] J. S. Colburn, M. A. Jensen, G. J. Pottie and Y Rahmat-Samii, UCLA Technical Report ENG-95-133, 1995.
[4] M. A. Jensen and Y. Rahmat-Samii, IEEE Proc., Jan. 1995.

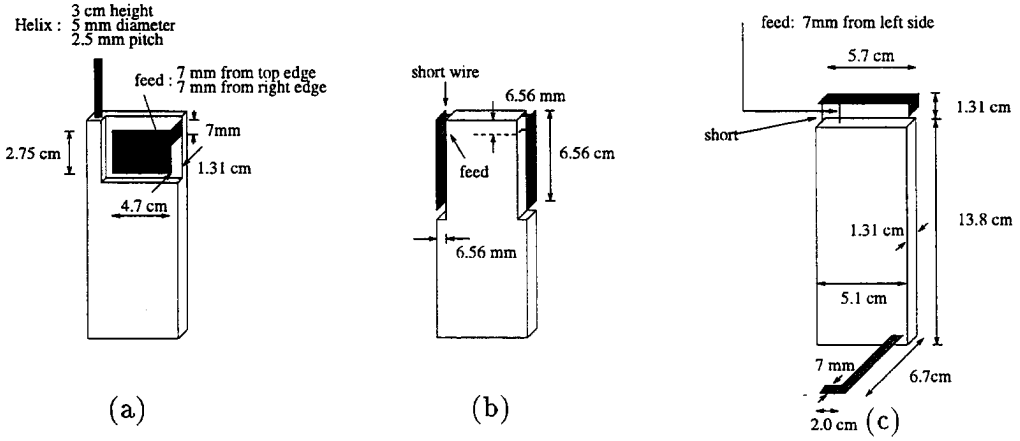


Figure 1: Handset antenna configurations considered: (a) top mounted helix and back mounted planar inverted F (PIFA); (b) side mounted PIFAs; (c) top mounted PIFA and "flip" monopole.

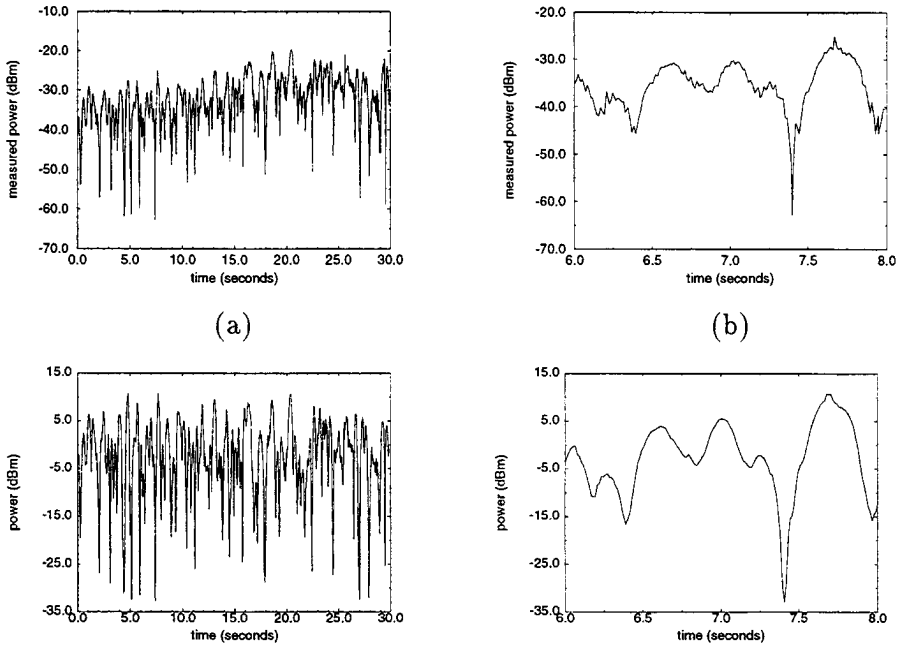


Figure 2: (a),(b) Recorded signal strength for one of the two dipole antennas spaced 0.06 λ apart. (c),(d) Processed data (2 second period demeaning and 3 point smoothing) of the raw data displayed (a) and (b).

Table 1: Power correlation coefficients calculated for the two side mounted PIFAs for path 1.

	Number of points in short term average				
time (s)	1	3	5	7	9
0.5	0.9714	0.9733	0.9752	0.9770	0.9786
1.0	0.5991	0.6009	0.6062	0.6155	0.6285
1.5	0.5422	0.5425	0.5460	0.5535	0.5643
2.0	0.5093	0.5100	0.5138	0.5213	0.5318
2.5	0.4934	0.4955	0.5012	0.5104	0.5225
3.0	0.4900	0.4924	0.4985	0.5082	0.5206
3.5	0.4980	0.5008	0.5071	0.5168	0.5292
4.0	0.4979	0.5003	0.5060	0.5152	0.5270
4.5	0.4848	0.4874	0.4932	0.5025	0.5145

Table 2: Power correlation coefficients for the proposed antenna configurations. NA denotes data not available.

Antenna Configuration	Path 1	Path 4	Path 5	Path 6
0.06 λ separated dipole	0.8039	0.8089	0.7362	0.7521
0.4 λ separated dipole	0.6143	0.6281	0.5739	0.6226
side PIFAs	0.5100	0.5460	0.6053	0.5735
back PIFA/helix	0.6117	0.6182	0.5961	NA
top PIFA/flip monopole	0.5385	0.6586	0.6109	NA

A CMOS Channel-Select Filter for a Direct-Conversion Wireless Receiver

P.J. Chang, A. Rofougaran, and A. Abidi

Proceedings of the IEEE Symposium on VLSI Circuits
pp. 62-63
June 1996

To be presented at the *Symposium on VLSI Circuits*, Honolulu, June 1996

A CMOS Channel-Select Filter for a Direct-Conversion Wireless Receiver

P. J. Chang, A. Rofougaran, and A. A. Abidi
Electrical Engineering Department
University of California
Los Angeles, CA 90095-1594

Introduction & Specifications

Whereas noise and nonlinearity in the RF front-end determine the *sensitivity* of a wireless receiver, the filters wholly govern its *selectivity*—that is, its ability to pick out a signal of interest in a frequency band, while rejecting all the other nearby signals populating that band. Filtering, like amplification, builds progressively in the receiver. The *RF preselect filter* attenuates the image to the desired channel, which would otherwise superimpose on the desired channel after downconversion, and it suppresses large out-of-band interferers which may overload the front-end. However, this filter's passband encompasses a block of frequencies surrounding the desired channel. A second filter at IF or baseband, with a much tighter passband on an absolute frequency scale, then selects the *one* desired channel.

In a direct-conversion or zero-IF receiver, this is a *lowpass* filter and may be integrated with the receiver. Some active channel-select filters have been described for use in direct-conversion receivers [1-4]. They are all bipolar continuous-time filters, which trade off large-signal handling against noise. A CMOS switched-capacitor (SC) filter, on the other hand, inherently decouples linearity from noise, but this has only been used once to date [5].

This paper describes the design and performance of a SC channel-select filter for a frequency-hopped spread-spectrum direct-conversion receiver operating in the 902-926 MHz ISM band (Fig.1). Data modulates a hopped carrier by one of two (or four) prescribed frequency offsets, such that the modulated spectrum is always contained within ±200 kHz of the carrier. Following preselection of the 26 MHz-wide ISM band at RF, an agile local oscillator synchronized to the hopping pattern of the received signal downconverts the band to bring the sought channel to DC. The channel-select filter passband extends 220 kHz around DC, and the adjacent channel 320 kHz away from the carrier defines the filter's stopband edge. System simulations suggest that if all ISM-band users in a cell abide by power control, then to detect the desired channel with a sufficiently high signal-to-interference ratio, the filter must attenuate adjacent channels by at least 45 dB.

Filter Architecture & Circuit Design

A 6th-order elliptic lowpass filter implements the sharp transition from 220 kHz to 300 kHz. FSK detection is insensitive to the non-constant group delay of this filter. The sampling rate of the SC filter is determined by the large stopband required. Assume, initially, that the RF preselect filter passes the ISM band, and has brickwall edges at the band edge. Then, if the desired channel hops to one edge of the ISM band, the channel-select filter must attenuate all other ISM users populating the 26 MHz to the other edge of the band. This implies a minimum sample rate of 26.22 MHz. However, as a typical 915 MHz dielectric preselect filter has a wide transition band, a 57.2 MHz sample rate is used in the SC filter to avoid aliasing near-ISM band interferers. This clock is conveniently obtained by dividing the 915 MHz local oscillator output by 16.

A *two-stage downsampling*, or decimation, architecture, reduces the capacitor spread in the filter (Fig.2). At the front-end, a simple 2nd-order Butterworth lowpass filter samples the continuous-time downconverted signal. The prefilter poles are selected to attenuate by 50 dB at 13 MHz. Next, a 6th-order elliptic filter comprising three biquads in cascade downsamples the prefilter output by 4. The passband of this stage is 220 kHz wide with 1 dB ripple, and the stopband, whose edge lies at 300 kHz, offers an attenuation of 50 dB. Thus, the prefilter suppresses the images of the elliptic filter passband appearing at multiples of the latter's 14.3 MHz clock.

The active filter also implements the baseband amplification in the receiver. With a voltage gain of 10× in the RF front-end [6] and a filter gain of 16×, the receiver can detect a 0.5μV input signal, which corresponds to a sensitivity of −116 dBm. This gain is distributed in the biquad stages according to their Q-factor (Fig.2) to scale the overall filter for maximum dynamic range.

The two main sources of noise in this filter are: the switch-induced *kT/C* noise at the prefilter input; and the referred input noise of the first biquad. This prototype was originally designed to also act as the downconversion mixer by subsampling the LNA output [7]. The track-mode bandwidth of the (passive) switch and capacitor network at the prefilter input is therefore greater than 1 GHz, and the small input capacitor results in a total input-referred noise in the passband of 70 nV/$\sqrt{\text{Hz}}$. The filter current is 4.6 mA. Since then, the filter has been redesigned to follow a continuous-time downconversion mixer [6], by making the prefilter sampling capacitor larger and allotting more gain to the prefilter. The referred input noise is thus lowered to 15 nV/$\sqrt{\text{Hz}}$, which is comparable to the output noise of a front-end with 3 dB noise figure and overall 10× conversion gain [6]. The current drain increases to 6.1 mA. This is still lower than the current extrapolated from previously published continuous-time bipolar channel-select filters, after they are scaled for the same noise level. For instance, a 7.5 kHz 6th-order elliptic lowpass filter for a paging receiver [1] will dissipate 8 mA after scaling, while a 172 kHz 3rd-order Butterworth filter will require 30 mA [3]. Neither filter's stopband has been evaluated to 50 MHz, nor are they capable of handling large-signals as well as this CMOS circuit.

Experimental Results & Discussion

The filter is implemented as a fully balanced circuit in a single-poly, double-metal 1μm CMOS process, and occupies an active area of 2.1×2.4 mm. A total of 200 pF is implemented on-chip as linear capacitors between gate poly and a heavy diffusion. The spread in capacitor value is 108:1. All transconductance op amps use single-cascode differential pairs. The circuit is evaluated at a 3.3V supply.

The measured filter response corresponds very well to the expected response (Fig.3), except for an additional 1 dB droop in the passband owing to incomplete settling in the high-gain biquad. Parasitic effects, such as on-chip signal feedthrough, limit the stopband attenuation to 65

dB on the large frequency scale up to 50 MHz (Fig.3(b)). The frequency response is unchanged when the front-end subsamples an input tone offset in frequency from 915 MHz, accomplishing both frequency-downconversion from RF and channel selection (Fig.3(a)). The measured input-referred noise spectrum is close to the expected value in the passband (Fig.4).

Of special interest in an active filter is its capability to handle large-signals, yet there are no standards on how to evaluate this. Therefore, the sources of distortion which matter to a particular receiver must be identified, and then they should be characterized in a way meaningful to the receiver architect. Harmonics of a single FSK-tone at 160 kHz this filter may produce are unimportant because they lie in the stopband. Similarly, intermodulation of a desired tone with a second tone in the passband is irrelevant, because such an interferer will cause receiver malfunction. However, certain combinations of *two tones in the stopband* produce intermodulation in the passband (Fig.5(a)). This is characterized by the extrapolated intercept between the filter output due to a single tone in the passband, and the 3rd-order intermodulation product of two tones in the stopband. The intercept point is lower by about 7 dB when subsampling 915 MHz. 2nd-order distortion in the filter front-end *detects amplitude-modulation* on a stopband signal, to produce spurious energy in the filter passband at DC (Fig.5(b)). This is similarly expressed as an input-referred intercept point. In both cases, the intercept point is higher than the 28 dBm output intercept point of the RF front-end [6], which means that the filter is not a bottleneck to receiver linearity.

Finally, DC offset accumulates through the non-autozeroed integrators used in the filter, and may saturate the subsequent limiting amplifier. The measured output offset ranges from 80 to 200 mV. Owing to a DC-biasing feedback which is capacitor decoupled off-chip [8], the limiting amplifier can accommodate up to ±100 mV input offset with little loss of dynamic range. A higher offset will require on-chip AC coupling of the signal.

1. D. W. H. Calder, "Audio Frequency Gyrator Filters for an Integrated Radio Paging Receiver," presented at *Int'l Conf on Mobile Radio Systems & Techniques*, York, UK, pp. 21-26, 1984.
2. J. F. Wilson, R. Youell, T. H. Richards, G. Luff, and R. Pilaski, "A Single-Chip VHF and UHF Receiver for Radio Paging," *IEEE J. of Solid State Circuits*, vol. 26, no. 12, pp. 1944-1950, 1991.
3. M. Koyama, T. Arai, H. Tanimoto, and Y. Yoshida, "A 2.5-V Active Low-Pass Filter Using All-n-p-n Gilbert Cells with a 1-Vp-p Linear Input Range," *IEEE J. of Solid-State Circuits*, vol. 28, no. 12, pp. 1246-1253, 1993.
4. C. Marshall, *et al.*, "A 2.7V GSM Transceiver ICs with On-Chip Filtering," presented at *Int'l Solid State Circuits Conf.*, San Francisco, pp. 148-149, 1995.
5. D. Haspeslagh, J. Ceuterick, L. Kiss, and J. Wenin, "BBTRX: A Baseband Transceiver for a Zero IF GSM Hand Portable Station," presented at *Custom IC Conf.*, San Diego, CA, pp. 10.7.1-10.7.4, 1992.
6. A. Rofougaran, J. Y.-C. Chang, M. Rofougaran, S. Khorram, and A. A. Abidi, "A 1 GHz CMOS RF Front-End IC with Wide Dynamic Range," presented at *European Solid-State Circuits Conf.*, Lille, France, pp. 250-253, 1995.
7. P. Y. Chan, A. Rofougaran, K. A. Ahmed, and A. A. Abidi, "A Highly Linear 1-GHz CMOS Downconversion Mixer," presented at *European Solid-State Circuits Conf.*, Sevilla, Spain, pp. 210-213, 1993.
8. S. Khorram, A. Rofougaran, and A. A. Abidi, "A CMOS Limiting Amplifier and Signal-Strength Indicator," presented at *Symp. on VLSI Circuits*, Kyoto, pp. 95-96, 1995.

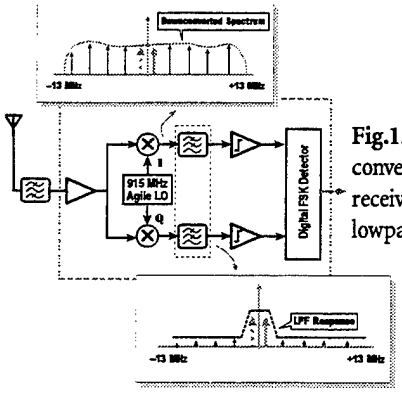


Fig.1. Block diagram of direct-conversion frequency-hopped FSK receiver. Role of channel-select lowpass filter is highlighted.

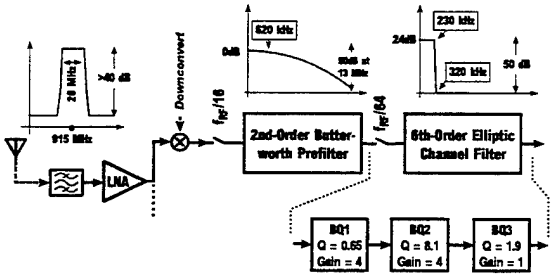


Fig.2. Per-channel filter architecture.

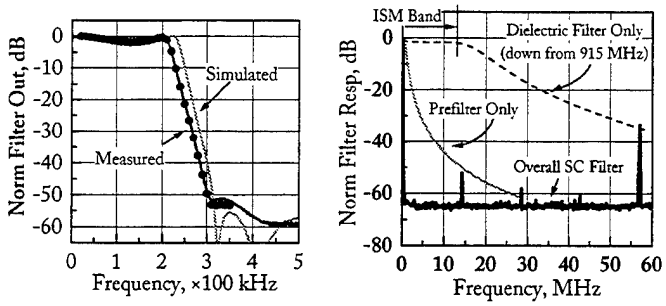


Fig.3. (a) Measured passband response; solid black line with baseband input, dots with 915 MHz input. (b) Large-scale stopband response.

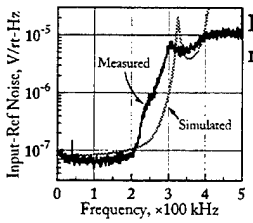


Fig.4. Measured input-referred noise of filter.

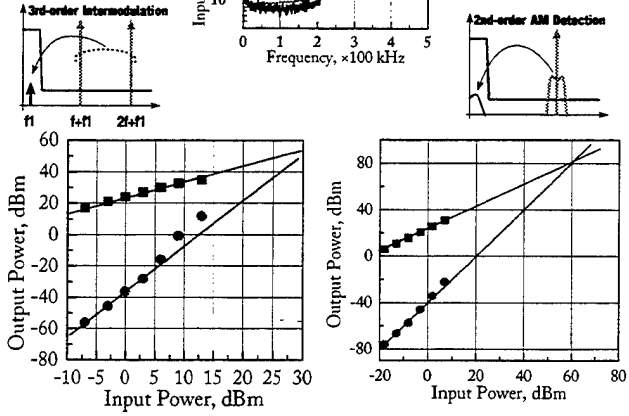


Fig.5. (a) 3rd-order intermodulation, and (b) 2nd-order distortion produce products in passband. Characterized by 3rd and 2nd-order intercepts.

Design Techniques for Silicon Compiler Implementations of High-Speed FIR Digital Filters

R.A. Hawley, B.C. Wong, T. Lin, J. Laskowski, and H. Samueli

IEEE Journal of Solid-State Circuits
vol. 31, no.5, pp. 656-667
May 1996

Design Techniques for Silicon Compiler Implementations of High-Speed FIR Digital Filters

Robert A. Hawley, *Student Member, IEEE*, Bennett C. Wong, *Member, IEEE*, Thu-ji Lin, *Member, IEEE*, Joe Laskowski, and Henry Samueli, *Member, IEEE*

Abstract—**Architecture design techniques for implementing both single-rate and multirate high throughput finite impulse response (FIR) digital filters are explored, with an emphasis on those which are applicable to automated integrated circuit layout techniques. Various parallel architectures are examined based on the criteria of achievable throughput versus hardware complexity. Well-known techniques for reduced complexity and computation time are briefly summarized, followed by the introduction of several new techniques which offer further gains in both throughput and circuitry reduction. An architecture for mirror-symmetric polyphase filter banks is derived which exploits the coefficient symmetry between multiple filters to reduce hardware. Finally, the evolution of a silicon compiler which utilizes all of these techniques is presented, and results are given for compiled filters along with comparisons to other compiled and custom FIR filter chips.**

I. Introduction

THE trend towards increasing data rates in digital signal processing systems has pushed the development and implementation of high-speed finite impulse response (FIR) digital filters beyond the capabilities of general-purpose programmable DSP chips, such as the Texas Instruments TMS320 series and the Motorola 56000 series. Dedicated programmable FIR filter chips are commercially available [1] which provide improved performance, allowing sample rates of up to 30 MHz and containing as many as eight taps on a single chip. An impressive high-speed programmable FIR chip was presented by Hatamian and Rao in [2] which operates at 100 MHz and contains 40 taps, but dissipates over 3 W of power due to its large complexity (240 000 transistors). High-speed and low-power applications require both increased parallelism and reduced complexity in order to allow both sampling rate and power dissipation goals to be met. A large number of programmable coefficient filter chips which utilize complexity reduction techniques have been published [3]–[6]. For many applications, reduced complexity may be achieved by eliminating programmability of the coefficients, thus allowing the hardware to be optimized for a particular fixed coefficient set. Coefficient recoding in forms such as canonic-signed digit (CSD) [7] and the use of alternative arithmetic systems such as the residue number system (RNS) [8] allow potential complexity reductions in either programmable or fixed filters. For fixed filter implementations, it is necessary to create custom silicon solutions for each application. The large number of applications for such application specific integrated circuits (ASIC's) would suggest that either a mask programmable solution or a compiled silicon solution [9]–[11] would be desirable. This paper will describe various techniques by which sufficient parallelism for high-speed operation may be achieved, while simultaneously constraining the solution to have a minimal complexity implementation for these structures. Most of these techniques are widely known and are briefly summarized as an introductory tutorial. Some new techniques which further reduce complexity have been developed and are presented in greater detail. In particular, a new technique which exploits the coefficient symmetries between subfilters in a polyphase filter bank is derived which offers similar hardware reductions as those found in linear phase filters. Silicon compilers which utilize these architectures and techniques have been developed and will be described subsequently. Finally, design examples for several compiled chips and hand-crafted chips are presented and compared.

Manuscript received November 1, 1993; revised October 9, 1995. This work was supported in part by research grants from the University of California MICRO Program, TRW Electronic Systems Group, Alcatel Network Systems, and the Advanced Research Projects Agency.

R. Hawley and H. Samueli are with Broadcom Corporation, Irvine, CA 92718, and the Integrated Circuits and Systems Laboratory, University of California, Los Angeles, CA 90024 USA.

J. Laskowski is with the Broadcom Corporation, Irvine, CA 92178 USA.

B. Wong and T. J. Lin are with PairGain Technologies, Tustin, CA 92680 USA.

Publisher Item Identifier S 0018-9200(96)03408-7.

II. FIR Filter Structures

The choice of structure for the implementation of an FIR filter includes consideration of factors such as hardware complexity and desired throughput. Many different structures exist, most of which provide some trade-off between complexity and throughput. For a dedicated application, the design choice then becomes the minimal complexity structure which can achieve a given throughput rate. The design choice for a silicon compiler, which must cover a range of applications, is not as straightforward. In fact, factors such as regularity and scalability (the ability to trade hardware for speed within the context of an architecture) become more important. Several structures have been investigated and the pros and cons of using each of these structures as the basis of a compiler are summarized briefly below.

A. Polyphase Structure

In the most general case of a filter, whether FIR or IIR, the input and output sampling rates need not be the same. Nonrational ratios of these sampling rates cannot be achieved in practice, but these may be approximated to an arbitrary precision by multirate filters with rationally ratioed sampling rates. The so-called polyphase structure may be utilized in order to achieve an efficient multirate filter implementation [12]. The rational sampling rate change can be achieved through the cascade of an interpolation and a decimation stage in which the input and output sampling rates differ by an integral number N. For this type of filter, an initial prototype filter is designed and then is divided into N subfilters by sampling the coefficients at N-spaced intervals. The resulting structures for interpolation-by-N and decimation-by-N are shown in Fig. 1(a) and (b), respectively. The polyphase structure results in an efficient use of the filter hardware since each subfilter operates at one-Nth of the required throughput rate. As a result, the implementation of polyphase multirate filters reduces to that of single-rate filters, so each subfilter can be realized using any of the many available single-rate digital filter architectures, several of which are described below.

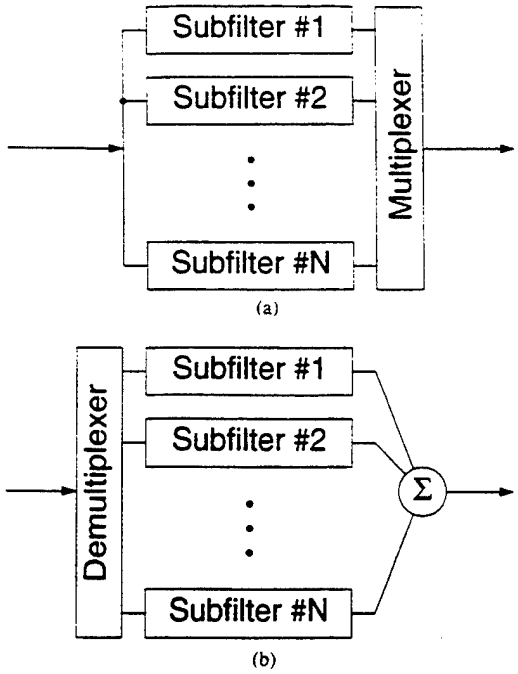


Fig. 1. (a) Interpolator structure and (b) decimator structure.

B. Direct Form Structure

A direct mapping of the difference equation describing an FIR filter into hardware results in the familiar direct form structure of a tapped delay line in which each of the delayed versions of the input is multiplied by the appropriate filter coefficient and the results are summed together to form the filter output. For a linear accumulation of the sum, the delay through an N-tap structure is $T_{mult} + (N - 1) \times T_{add}$. Of course, the accumulation can instead be performed by a tree of adders as suggested by Reutz in [13], resulting in a delay of $T_{mult} + (\log_{1.5} N) \times T_{add}$. The direct form structure with linear accumulation is very amenable to implementation in a compiler because of its high degree of regularity. Each tap of the filter is equivalent, so arbitrary length filters can be generated by simply cascading filter taps. Unfortunately, the delay through the filter increases linearly with the number of taps and thus may not be a good candidate for high-speed operation. The direct form structure with a tree of adders would provide superior performance, since its delay increases logarithmically with increasing filter length. Additionally, the adder tree could be efficiently pipelined to eliminate this dependence entirely. The possible types of accumulation trees mirror those in multipliers, including such structures as the Wallace and Dadda trees. Unfortunately, these trees tend to be much more irregular than the linear accumulation structure and thus will be more difficult to implement with a highly structured compiler, such as those described below. Ignoring the effect of routing on the circuit area, both the linear and tree accumulation structures should occupy roughly equivalent areas since they each contain approximately N multipliers, $(N - 1) \times WL_{int}$ single-bit adders, and $(N - 1) \times WL_{ext}$ single-bit registers, where WL_{int} is the internal wordlength and WL_{ext} is the input wordlength.

C. Transpose Direct Form Structure

Application of the transposition theorem results in the transpose direct form structure in which the input is fed to all of the coefficient multipliers in parallel and the results are accumulated over N sample periods. This structure retains the regularity of the linear accumulation direct form structure but has a delay equal to $T_{mult} + T_{add}$ and is, therefore, independent of the filter length. One of the primary disadvantages of this structure is that the registers occur in the accumulation path and therefore must have the length of the internal wordlength of the filter rather than the possibly shorter input wordlength. Another disadvantage is the large loading on the input data broadcast bus since all multipliers are fed in parallel. These are relatively minor problems, however, which may be partially reduced through other techniques. Ulbrich and Noll have proposed this architecture to retain the full advantage of bit-parallel processing [14].

D. Linear Phase Direct Form Structures

In many filtering applications phase distortion cannot be tolerated, and thus the filters are required to have a linear phase response. It is well known that an FIR filter can be guaranteed to have an exact linear phase response if the coefficients are either symmetric or antisymmetric about the center tap. This coefficient symmetry can then be exploited by sharing the multipliers between the (anti-)symmetric taps. Of course, this symmetry feature exists in both the direct form and transpose form struc-

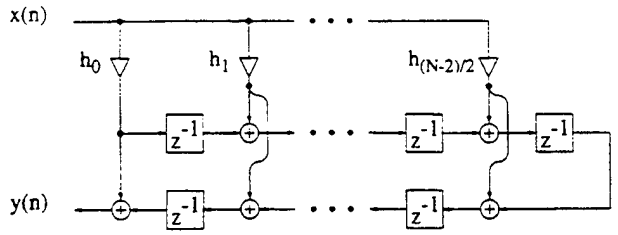


Fig. 2. Linear phase transpose direct form filter structure.

tures, with the same architectural features as in the nonlinear phase case. The only differences are the reduction in the number of multipliers by about one-half along with reduced coefficient routing and a slight increase in some signal routing due to the shared multipliers. An example of this structure for an even order transpose direct form filter is shown in Fig. 2.

E. Bit Plane Structure

An important issue which has not been touched upon in the discussion thus far has been the issue of the required internal wordlength of the filter. The use of finite wordlength arithmetic in the filter introduces truncation and/or rounding errors which deteriorate the signal-to-noise ratio at the filter output. To minimize these effects, guard bits are usually added to the right of the least significant bit (LSB) of the input data to determine the internal wordlength. The actual number of guard bits required is dependent on the number of filter taps, the composition of the coefficients, the desired signal-to-noise ratio (SNR) at the filter output, and the input data wordlength. Typically 2–6 guard bits are sufficient for most applications.

The bit-plane architecture [15] can avoid these LSB guard bits because of the unique way by which filter computations are carried out. Fig. 3 shows the bit-plane realization for a transpose form four-tap FIR filter. In the first bit-plane, the least significant partial products of all of the coefficients are computed and accumulated. The output of the first bit plane is then shifted to the right by one bit, thus truncating the LSB, and delivered to the second bit plane. Four registers are required along the input data line between the two adjacent bit-planes for the purpose of deskewing the data. Subsequently, the second least significant partial products are processed, etc. The lower significant bits of the partial sums are always generated earlier, so that they can be discarded without effecting the higher significant bits in the final output. Hence, as far as output SNR is concerned, no LSB guard bits are required to obtain the full precision output in the bit-plane structure. It is also apparent that bit-level pipelining is inherent in the filter structure; hence, it may be a good candidate for high-speed applications. Additionally, the structure is quite regular, so implementation as a compiler should not prove particularly difficult.

Unfortunately, the bit-plane structure possesses several undesirable qualities. The first is that guard bits should be provided to the left of the most significant bit (MSB) of the input data to accommodate potential overflows. In

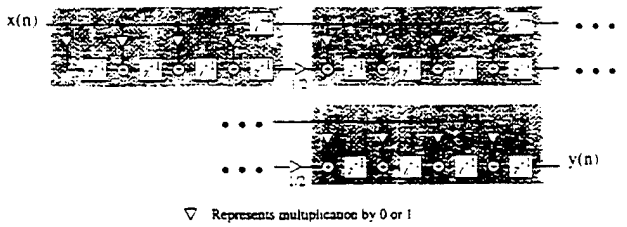


Fig. 3. Bit-plane structure for a four-tap FIR filter.

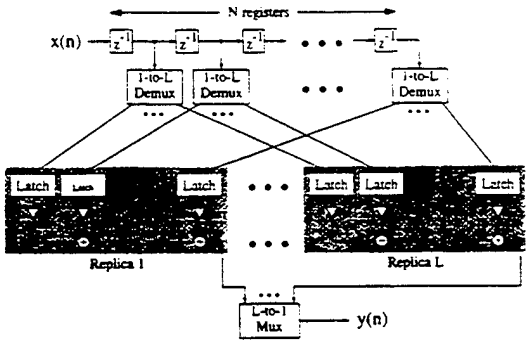


Fig. 4. Order-L block processed realization of N-tap direct form FIR filter.

general, for an N-tap filter, $\log_2 N$ MSB guard bits are required to prevent overflows. A second disadvantage has to do with the number of registers which are required for filters with many taps. Referring to Fig. 3, it can be seen that the total number of registers required is equal to

$$2 \times WL_{\text{int}} \times N \times WL_{\text{ext}}. \quad (1)$$

The number of adders required will depend on the number of nonzero bits in the coefficients, but will be upper bounded by (1). A final disadvantage of this architecture is the latency which is inherent in the structure. For an N-tap filter, this latency will be $N \times WL_{\text{ext}}$.

F. Block Processing Structures

Pipelining of an FIR filter increases its throughput by partitioning an operation into smaller chunks and then overlapping the operations. Another way to increase the throughput is to provide replicas of the hardware in which a block of input data is sent to each replica in turn [16]. An output is then generated by each replica in a round-robin fashion. Fig. 4 shows an order-L block processed realization of an N-tap direct form FIR filter. A similar structure exists for transpose direct form filters. Thus, by using each of the filters to produce only $1/L$th of the output samples, the throughput can be theoretically increased by the factor L. In practice, the delays through the multiplexers and demultiplexers as well as second-order loading effects limit the throughput increase to somewhat less than L and also limit the maximum value of L which can be effectively used.

G. Comparison of the Structures

The choice of structure for compiler implementation reduces to the common speed versus area trade-off with the

TABLE I
FIR STRUCTURE COMPARISONS

FIR Structure	Delay	Active Area
Direct Form Linear Accumulation	$T_{mult} + N \times T_{add}$	$(N-1) \times [WL_{int} \times A_{add} + WL_{ext} \times A_{reg}] + N \times A_{mult}$
Direct Form Tree Accumulation	$T_{mult} + \log_{1.5} N \times T_{add}$	$(N-1) \times [WL_{int} \times A_{add} + WL_{ext} \times A_{reg}] + N \times A_{mult}$
Transpose Direct Form	$T_{mult} + T_{add}$	$(N-1) \times [WL_{int} \times A_{add} + WL_{int} \times A_{reg}] + N \times A_{mult}$
Bit-Plane Structure	T_{add}	$2 \times N \times WL_{int} \times WL_{ext} \times (A_{add} + A_{reg})$

qualification that the chosen structure should not be so irregular so as to preclude efficient compiler based design. Table I summarizes the relative delays and areas for the non block processed structures. Block processing can increase the speed of any of these structures by a large factor, but will also incur an area increase by the same factor. This should be avoided if at all possible, and one of the other structures should be used. The bit-plane structure seems to offer the ultimate in speed if the time T_{mult} is much greater than T_{add}, but the area penalty may be excessive unless the area of the multipliers is large. If these delays are comparable, or if the filter can be pipelined, the transpose direct form would seem to offer the best speed versus area trade-off, especially when reduced complexity multipliers can be used. This was seen to be the case in the implementation of some custom chips which will be described later, so the transpose direct form was used as the basic architecture for the filter compilers.

III. Architectural Techniques for High-Speed Operation

A. Coefficient Recoding

The primary limitation in the achievable throughput in the transpose direct form structure is the multiplier delay. For the case of fixed-coefficient filters, which we are considering here, there are several options for reducing the complexity of the multipliers and increasing their throughput. Since dedicated multipliers can be used, the multiplication operation is reduced to a series of hard wired shifts and additions. The delay through this structure is equal to $(L - 1) \times T_{add}$, where L is the number of nonzero bits in the coefficient representation. It makes sense then to use an efficient coding of the coefficients such as radix-2 signed-digit coding, which is made up of the redundant digit set $(-1, 0, 1)$ [17]. The radix-2 signed-digit representation of a fractional number x has the general form

$$x = \sum_{k=1}^{L} s_k 2^{-p_k} \tag{2}$$

where $s_k \in \{-1, 0, 1\}$ and $p_k \in \{0, 1, \cdots, M\}$. This representation has at most L nonzero digits within an $M + 1$ digit wordlength. The CSD code is defined to be a minimal representation which achieves the least number of nonzero digits for which no two nonzero digits are adjacent. The distribution of the CSD codes for different values of M and L is nonuniform, but it has been shown that with proper design, filters using such codes may be generated with little degradation in the desired frequency domain characteristics. In fact, only one nonzero CSD digit is required for approximately each 20 dB of stopband attenuation [18].

B. Redundant Addition Schemes

The next limiting factor in the achievable throughput is due to the classic problem of carry propagation. The simple carry-ripple adder circuit is limited by its ripple path delay which is proportional to the wordlength of the adder, or $O(N)$. Because this delay is often not acceptable, there are numerous techniques available to speed up carry propagation. These range from the relatively straightforward carry-lookahead, carry-select, and carry-skip techniques [17] to the more sophisticated conditional-sum approach [19]. Each of these approaches attempts to reduce the delay of the carry propagation chain to below $O(N)$, but none achieves a smaller delay than $O(\log_2 N)$ [17], and these often come at the expense of a great increase in hardware complexity.

Another approach to the carry propagate problem is to remove it completely through a redundant digit coding scheme such as signed-digit addition or carry-save addition. In carry-save addition, both a sum and carry bit are kept for each bit position in the word, thus presenting a redundant coding scheme. This technique may be used to sum an arbitrary number of words, but the final result will always contain two bits per bit position. These must then be decoded into binary digits through the use of some type of carry propagate adder. Carry-save addition may then be used to accumulate partial sums within the core of a filter, but a carry-propagate adder will still be required at the filter output to decode the carry-save result. Since only one carry-propagate adder is required, it is possible to use one of the more complex fast adder techniques without an excessive impact on the total area. There are a few drawbacks to the carry-save scheme, however, with the most important of these being the negative impact of doubling the number of registers required within the filter core. This increase in area is a price that must be paid in order to achieve a high throughput which is the goal of this work.

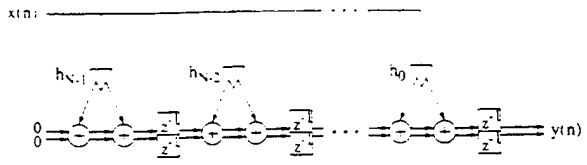


Fig. 5. Modified transpose direct form FIR filter structure.

These changes to the transpose direct form structure are shown in Fig. 5, in which the multipliers are now carry-save CSD multipliers.

C. Pipelining

With the adoption of CSD multipliers and carry-save addition, the critical path of the filter has been greatly reduced, but can be reduced further through pipelining of the structure. It is quite evident that both the transpose direct form structure and the folded transpose direct form structure can be easily pipelined anywhere within the multipliers since they form a simple cut-set. It will prove more difficult, however, to pipeline between the tap adders.

In order to study the pipelining required, it is necessary to know the number of nonzero bits in the CSD coefficient codes. The number of bits required will depend on the required stopband attenuation of the filters to be generated. As 60–80 dB of stopband attenuation is sufficient for most filtering applications, no more than four nonzero CSD bits should be required [18]. For the case of four nonzero bits, the CSD multiplier reduces to two adders to perform the required 4-to-2 vector reduction. Pipelining to a two-adder delay can then be achieved by placing pipeline registers between the multipliers and the adders as shown in Fig. 6(a). It is easily seen from Fig. 6(a) that 2-, 3-, and 4-bit CSD coefficients will require two pipeline registers per filter tap, while 1-bit CSD coefficients will require one pipeline register and 0-bit coefficients will not require any pipeline registers to achieve a two-adder delay critical path. Pipelining to a single adder delay as in Fig. 6(b) will require substantially more hardware and will deliver diminishing benefits since register set-up and clock-to-output times and the delay of the data broadcast will now form a significant portion of the cycle time. The register cost per filter tap for bit-level pipelining is $3 \times (L - 1)$ pipeline registers for an L-bit CSD coefficient. Because of this heavy cost and diminished returns, only pipelining to the two-adder delay level will be considered.

D. Sign Extension Elimination

With speed limitations in the multipliers and adders thus reduced, other factors begin to gain importance in determining the final operating speed of the generated filters. The most critical of these is the delay due to data broadcast. Because of the nature of the transpose structure, there will necessarily be a large load on the input data bus which cannot be easily eliminated. Some work has exploited the common factors in the coefficient set to produce a nested

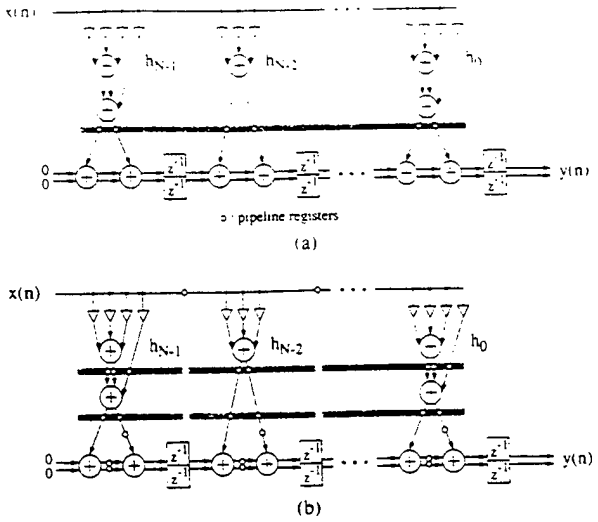


Fig. 6. (a) Transpose direct form filter with two-adder delay pipelining and (b) transpose direct form filter with one-adder delay pipelining.

generation of the data-coefficient products. While this may simplify the generation of products and reduce the loading on the data broadcast lines, it is coefficient dependent and extremely irregular. Although this can provide significant reductions in complexity and loading for some filters, it is nearly impossible to implement in the structured compilers described below and so is not described further.

The problem of excessive loading is compounded by the need to sign-extend the 2's complement numbers in the multipliers. Traditionally this is done by appending the sign digit to the left of the other digits as they are right-shifted. For example, suppose that we wish to multiply the fractional 8-bit input data word

$$b_0.b_1b_2b_3b_4b_5b_6b_7 \tag{3}$$

by the CSD coefficient 0.015625 or 2^{-6}. The multiplication is accomplished by right-shifting the data word by 6-bit positions while sign-extending the MSB

$$b_0.b_0b_0b_0b_0b_0b_0b_1b_2b_3b_4b_5b_6b_7. \tag{4}$$

This will lead to a far greater loading on the MSB as compared to the other bits on the input data bus, which will cause a myriad of layout problems. The buffers for the MSB will need to be much larger, causing additional power supply inductive drops, increased power consumption, and wider power and data broadcast buses. This will destroy the regularity of the filters, thereby complicating the filter design and leading to an increase in area. There are other ways to perform sign extension, however, which will allow these problems to be eliminated. Borrowing a trick from 2's complement multiplier theory, we can represent the shifted data word in (4) as the sum of the two vectors

$$\begin{aligned} &0.000000\bar{b}_0b_1b_2b_3b_4b_5b_6b_7 \\ &+\ 1.1111111 \end{aligned} \tag{5}$$

TABLE II
NUMBER OF LINEAR PHASE SUBFILTERS IF PROTOTYPE FILTER IS LINEAR PHASE

Filter Length	Sampling Rate Conversion Ratio	# of Linear Phase Subfilters
Even	Even	0
Even	Odd	1
Odd	Odd	1
Odd	Even	2

where the overbar denotes a bit complement. Thus, the MSB sign extension has been eliminated and replaced by an all-ones "compensation vector" which must be added to the filter output. These constant ones vectors generated by each shift in the CSD multipliers can be collected together and presummed to form a *global* compensation vector, so that only a single addition will be required to incorporate the compensation vectors into the final accumulated sum. Referring back to Fig. 5, we see that there are two unused adder inputs in the first tap of the filter, so one of these may be used to insert the global compensation vector. Therefore, the sign extension problem can be completely eliminated without any hardware overhead.

IV. New Techniques for Further Hardware Reduction

Many of the previously mentioned architectural techniques for reduced complexity filters are well known and have been used in other silicon compilers. This section will introduce some new techniques which can further reduce filter complexity and improve performance. These have been incorporated to varying degrees in the compilers which will be presented in Section V.

A. Mirror Symmetric Filter Pairs

When dealing with multirate filters, the polyphase structure is usually chosen for its reduced complexity and improved performance. When the prototype filter is linear phase, however, the decomposition of the filter into L subfilters will usually result in nonlinear phase subfilters, and, thus, possibly increased complexity as compared to the prototype filter. Since the decomposition into subfilters is accomplished by sampling every Lth coefficient of the original impulse response, those subfilters resulting from sampling which is symmetric about the center tap will be linear phase, while the other subfilters will not. In fact, at most, two of the subfilters will be linear phase as is summarized in Table II. Thus, the remaining nonlinear phase subfilters cannot use the folded structure and will require a larger number of multipliers to implement.

Example 1: Suppose an interpolate-by-3 function is to be implemented with a 50-tap linear phase prototype filter with symmetric coefficients, using a polyphase structure. Because of the symmetry of the coefficients [i.e., $h(n) = h(49 - n)$], the difference equation can be written in the form

$$y(n) = \sum_{k=0}^{49} h(k)x(n-k)$$
$$= \sum_{k=0}^{24} h(k)[x(n-k) + x(n+k-49)] \quad (6)$$

and thus only 25 multipliers are required. After implementation in the polyphase structure of Fig. 1(b), the difference equations of the three subfilters will be

$$y_0(n) = \sum_{k=0}^{16} h(3k)x(n-3k)$$
$$y_1(n) = \sum_{k=0}^{16} h(3k+1)x(n-3k-1)$$
$$y_2(n) = \sum_{k=0}^{15} h(3k+2)x(n-3k-2)$$
$$= \sum_{k=0}^{7} h(3k+2)[x(n-3k-2) + x(n+3k-47)]. \quad (7)$$

Because only the third filter is linear phase, the resulting polyphase filter will require $17 + 17 + 8 = 42$ multipliers to implement. This is still fewer multipliers than would be required if the prototype filter was not linear phase, however.

It is interesting to examine the form of the nonlinear phase subfilters which are derived from the linear phase prototype filter. Examined individually, each filter has completely asymmetric coefficients, so no reduced form structure is possible. If the filters are examined together, however, some symmetries between the filters can be noted. In particular, the filters can be grouped into pairs in which both filters in each pair have the same set of coefficients; however, the coefficients occur in reverse order. Because of this property, we name the filters *mirror symmetric filter pairs*. This property can be exploited to reduce the number of multipliers required to the same as in the original prototype filter, while maintaining the polyphase structure, with its desirable properties.

Example 2: Continuing from Example 1, we recast the difference equations of the first two subfilters in the fol-

lowing form:

$$y_0(n) = \sum_{k=0}^{16} h(3k)x(n - 3k) = \sum_{k=0}^{16} h_0(k)x_0(n - k)$$

$$y_1(n) = \sum_{k=0}^{16} h(3k + 1)x(n - 3k - 1)$$

$$= \sum_{k=0}^{16} h_1(k)x_1(n - k)$$

$$= \sum_{k=0}^{16} h_0(16 - k)x_1(n - k) \quad (8)$$

where

$$x_0(n - k) = x(n - 3k)$$

$$x_1(n - k) = x(n - 3k - 1)$$

$$h_0(k) = h(3k) = h(49 - 3k)$$

$$h_1(k) = h(3k + 1) = h(48 - 3k)$$

$$= h(3(16 - k)) = h_0(16 - k). \quad (9)$$

Thus, because $h_1(k) = h_0(16 - k)$, we say that the filters form a mirror symmetric pair.

This fact would be interesting, but not very useful, if the filters were completely independent, sharing neither inputs nor outputs. This is not the case, however, for the polyphase filter structures, and therefore multipliers can be shared between the filters in a mirror symmetric pair, resulting in lower complexity structures.

1) Interpolator Structures: The subfilters in a polyphase interpolator share a common input, which makes it a simple matter to derive a direct form transpose structure for the mirror symmetric filter pair. This is illustrated in Fig. 7(a) which shares the multipliers between the two filters in the mirror symmetric pair and simply accumulates the results in opposite directions in order to form the correct transfer functions. Sharing the multipliers in a direct form structure is not as straightforward, however, and an $(N - 1)$ pipeline latency must be introduced to make the structure physically realizable. As shown in Fig. 7(b), for a direct form structure, the penalty paid in order to share the multipliers is $3N - 3$ additional registers.

Example 3: Continuing from Example 2, we note that filters $h_0(n)$ and $h_1(n)$ can be implemented as a mirror symmetric pair because

$$y_0(n) = \sum_{k=0}^{16} h_0(k)x_0(n - k)$$

$$y_1(n) = \sum_{k=0}^{16} h_0(16 - k)x_0(n - k). \quad (10)$$

Thus, the products $h_0(k)x_0(n - k)$ only need to be calculated once for the pair of filters, reducing the number of multipliers required for the overall polyphase structure from 42 to 25. Thus, by exploiting the mirror symmetric filter pairs, the full gain of using a linear phase prototype filter is achieved, along with the more efficient processing of the polyphase structure.

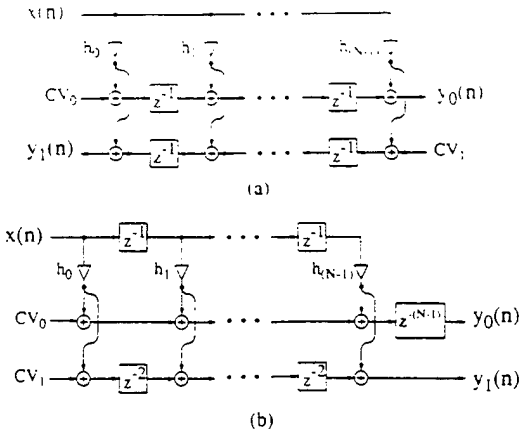


Fig. 7. (a) Transpose direct form mirror symmetric filter pair for interpolation and (b) direct form mirror symmetric filter pair for interpolation.

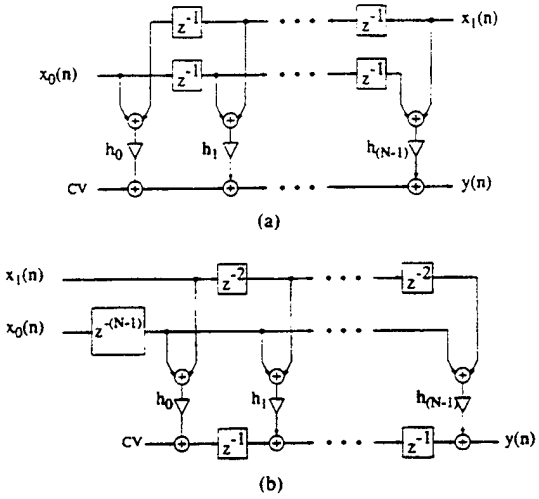


Fig. 8. (a) Direct form mirror symmetric filter pair for decimation and (b) transpose direct form mirror symmetric filter pair for decimation.

2) Decimator Structures: In contrast to interpolation, decimation filter pairs have different inputs and their outputs are summed together. In order to allow the coefficient symmetry to be exploited, the output adders are pushed through the filter structure to the inputs of the multipliers, thus allowing the multipliers to be shared. This is illustrated in Fig. 8(a) and (b). In the case of decimators, using the mirror symmetric property of the filter incurs no penalties for the direct form structure but costs $3N - 3$ additional registers for the transpose direct form structure.

In designing multirate polyphase filters, the usual practice is to use the direct form for interpolators and the transpose direct form for decimators since these allow the registers to be shared between the filters. If the prototype filter is linear phase, however, superior results may be obtained by reversing the choice of structures, thereby allowing multipliers to be shared between the filters in each mirror symmetric pair without an additional register penalty.

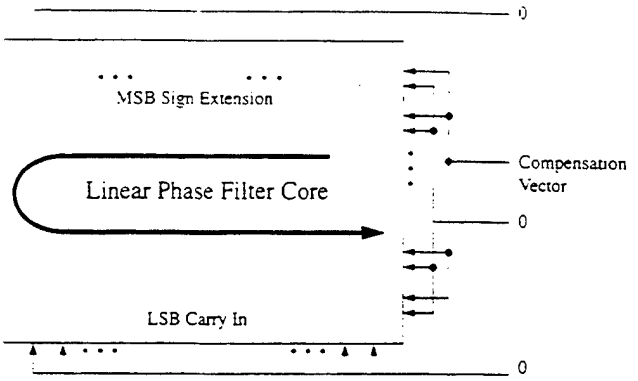


Fig. 9. Signals in a linear phase FIR filter which are known *a priori*.

B. Automated Circuitry Elimination

The fully parallel nature of these types of FIR filters makes them conceptually simple to implement as a *filter order* × *wordlength* array of identical cells, with known inputs shorted to Vdd or Gnd as required. This is very inefficient since it does not attempt to exploit any of the features of the coefficients, their positions within the filter, or the signal values which are known at the time of filter compilation. The compilers presented below make use of this knowledge to varying degrees in order to arrive at an efficient solution. The first technique is quite obvious and makes use of the knowledge of the number of CSD bits in each coefficient to eliminate unneeded circuitry. Thus, the single-bit cell compiler generates cells with different floorplans depending on the number of nonzero CSD bits in each coefficient. The second technique, which is more subtle, examines the inputs of each bit cell which are known *a priori* in order to eliminate unnecessary circuitry. This knowledge of certain filter input values comes from three primary sources as is shown graphically in Fig. 9. First, one of the data input buses to the first tap of the filter is tied to logic zero while the other input is set equal to the compensation vector previously described. The second known inputs are the zeros tied to the MSB's of the various taps as part of the sign-extension scheme and zeros tied to the LSB's of the taps in which the input data bus is not shifted beyond the internal wordlength of the filter. Finally, all of the carry inputs to the LSB's in each tap are known to be zero. To make use of this knowledge, the compiler propagates this information throughout the filter core based on addition rules and a knowledge of the signal interconnect to determine all of the known nodes in the filter to be implemented. The bit cell compiler then uses this information to only place circuitry which is connected to unknown inputs in each bit cell position. This can result in the elimination of a large number of transistors in the filter without affecting the overall filter floorplan, thereby reducing the loading on many signal lines. This in turn results in improved speed and power dissipation, but cannot typically be exploited to save area due to the somewhat random nature of the positions of the eliminated circuitry. Quantifying the effects of this circuitry elimination are not easy, but some simple calculations may be made. Several example filters have been compiled with the average transistor count reduction on the order of 25–30%. Of these, perhaps 50% are connected to nodes which would never switch and thus do not contribute to the circuit delay or power dissipation. If we assume that transistor loading and parasitic wiring loading form roughly equal shares of the total chip capacitance, this circuitry elimination technique can save about 7.5% of the total power dissipation of the chip. In addition, the yield is enhanced since the active area is reduced.

V. Silicon Compilers for FIR Filters

Because of the high-speed requirements for the filters presented, typically 200 MHz or greater, a silicon compilation technique which enforces strict floorplanning rules is utilized. This choice rules out the more typical design practice of creating a compiler which is based on logic synthesis techniques for netlist creation and standard cell place and route for layout implementation. A brief description of some silicon compilers for FIR filters written at UCLA are described below.

A. FIRGEN

FIRGEN was the original fully parallel FIR filter compiler written at UCLA. It uses a nonpipelined transpose direct form structure with CSD multipliers and carry-save adders. It does not exploit the symmetry in linear phase filters nor does it generate any of the special multirate structures described previously. An in-depth look at the specifics of its architecture has been published in [20] and so is not included here. Some filters generated by FIRGEN are included in Section VI for comparison. A high-level filter design interface was added to FIRGEN to ease the task of designing filters, generating and optimizing CSD coefficients, and linking to layout generation. Compilers which were designed subsequently also make use of this high-level interface.

B. LPFIR

In order to take advantage of the coefficient symmetry in linear phase filters, the compiler LPFIR was written. It shares the basic architecture used by FIRGEN but uses the folded structure, circuitry reduction techniques, and the improved antisymmetric coefficient handling described previously. Details on LPFIR may be found in [21]. Both FIRGEN and LPFIR are based on the LAGER IV silicon compilation system developed by University of California, Berkeley and are now part of the general release of the tools [22].

C. GenFIR

The most recent FIR filter compiler written at UCLA, GenFIR, is intended to be a general purpose compiler for many types of FIR filters. It can generate single-rate or

multirate filters with either transpose direct form or folded transpose direct form structures and also exploits the mirror symmetric properties of subfilter pairs in polyphase multirate filters in order to achieve more efficient layouts. The compiler uses all of the architectural and circuit reduction techniques described previously and incorporates two important additional features: design rule independence and automated device sizing, which will now be described.

1) Design Rule Independence: GenFIR is written in the Genie language which is a part of the GDT Designer Tools from Mentor Graphics. This makes it substantially different from the previous two compilers in several ways. Most compilers work by tiling together preexisting cells from a library which are created by the traditional technique of drawing polygons representing circuit elements on a CRT. The use of the compiler under a different IC processing technology, other than that for which the cells were originally designed, requires the creation of a new cell library dedicated to the new fabrication process. Conversely, GDT continues the compiled approach down to the cell level. That is, layout is performed by procedural programs, called layout generators, which implement the circuit layout in a symbolic manner, through the use of technology independent spacing variables and calls to high-level tools such as an incremental compactor and several routers [23]. Thus, by running these design rule independent generators with a given target technology, a library of process specific cells can be compiled. Layouts have been generated in technologies ranging from 2.0-μm MOSIS scalable CMOS to TRW's 0.5-μm radiation-hardened CMOS process using the same set of layout generators. This methodology is continued up through the design hierarchy, resulting in a compiler which is largely independent of specific design rules.

2) Automated Device Sizing: In addition to the ability to produce design rule error-free layout, a compiler which is technology independent also requires the ability to produce layouts with the correct electrical characteristics for the target process. The technology files in GDT contain both layout and electrical information about the IC processes, which the compiler extracts and uses to calculate the optimal sizes for the transistors, signal lines, and power lines in the different layout cells. For example, the load on a clock signal line due to both the transistors being driven and the parasitic loading of the wires is calculated by the compiler, and this information is fed to a buffer generator which generates the appropriate buffer tree to drive this load. Electromigration calculations are also automatically performed to determine the required width of the signal lines to meet process limitations. Any of these calculated values may be overridden by the designer, however, if desired.

VI. Comparisons of VLSI Implementations

The techniques presented in this paper have evolved gradually over time and have been incorporated in various handcrafted and compiled chips developed at UCLA. A summary of these chips is presented in Table III. The first three entries represent full custom handcrafted designs which were developed without the use of compiler tools. The final three entries were each produced by one of the compilers described in Section V. All of these chips were fabricated and tested and are fully functional at the speed given in Table III. A direct comparison of these chips to each other is difficult for several reasons. Different numbers of taps and different wordlengths can be normalized out of the figures of interest without too much difficulty. It proves more difficult, however, to remove the dependence on processing technology, numbers of nonzero CSD bits in the coefficients, and degree of pipelining from any figure of merit which is derived to rate the quality of the chip or compiler. Instead, the FIR filters contained in the chips listed in Table III were recompiled with GenFIR in the same technology in order to obtain more reasonable comparisons. These results are summarized in Table IV. The areas compared are the areas of the filter cores only, in order to allow a more direct comparison between the filters without the additional hardware for the control circuitry and the pad rings. Speed results for the GenFIR compiled filters should be similar to the originals since the same architecture is used in both cases. No comparison is made with the first chip in Table III since its architecture is much different from the one used by the compilers. There are two important conclusions which can be drawn from these comparisons. The first is that the compiled chips are quite close in area to the full custom handcrafted chips, and the second is that the filters compiled by GenFIR are substantially smaller than those produced by the earlier compilers, demonstrating the further improvements which have been obtained in both complexity reduction and compiler efficiency.

Example 4: As a final example, a filter was compiled by GenFIR to discover how many taps could be placed on an area of 1 cm^2 in a 1.0-μm CMOS technology. The chosen design is an interpolate-by-8 filter based on a linear phase prototype filter. This results in four pairs of mirror symmetric filters when the polyphase structure is used. The overhead of pads and the input and output control circuitry required approximately 0.2 cm^2 of area, leaving 0.8 cm^2 for the filters themselves. It was found that a 320-tap prototype filter with a 16-bit internal wordlength could be generated in the allowed area, resulting in a filter core of dimensions 1.2 cm $\times$ 0.67 cm. Thus, the tap density for these filters is approximately 400-taps/cm^2. The average number of CSD bits per tap coefficient is 2.28, and the total transistor count of the 320-tap filter core is 313 000. The overall chip dimensions including pads are 1.3 cm $\times$ 0.75 cm. The layout plot of the filter is shown in Fig. 10. Because the use of the polyphase structure allows the filters to operate at one-eighth of the clock frequency, the effective throughput of this chip will inevitably be limited by the multiplexer at the chip output. Because of this, the filters were compiled without the pipelining option in order to save area. Without pipelining,

TABLE III
LIST OF CUSTOM AND COMPILED FILTER CHIPS

Filter Structure	Compiler	Technology	Transistors	Area	Speed
X/SIN X Predistortion Filter [24] 11 Taps, 16-bit Wordlength Non-folded Structure Bit-level Pipelined	Hand Layout	1.0-μm Rad Hard CMOS	14,000	3.45 mm × 4.25 mm	200 MHz
QAM Modulator [25] 40 Taps, 14-bit Wordlength Interpolate by 4 Polyphase Structure with Non-folded Subfilters	Hand Layout	1.2-μm CMOS	46,000	7.01 mm × 4.87 mm	200 MHz
QAM Demodulator [25] 40 Taps, 16-bit Wordlength Decimate by 2 Polyphase Structure with Non-folded Subfilters	Hand Layout	1.2-μm CMOS	57,000	7.69 mm × 4.69 mm	200 MHz
X/SIN X Predistortion Filter [20] 11 Taps, 14-bit Wordlength Non-folded Structure	FIRGEN	1.2-μm CMOS	12,500	4.52 mm × 2.38 mm	112 MHz
Half-Band Filter [21] 43 Taps, 20-bit Wordlength Linear Phase Folded Structure	LPFIR	1.2-μm CMOS	66,000	6.5 mm × 6.3 mm	150 MHz
Hilbert Transform Filter [26] 43 Taps, 18-bit Wordlength Decimate/Interpolate by 2 Polyphase Structure with Linear Phase Folded Subfilters	GenFIR	1.0-μm CMOS	45,000	7.07 mm × 3.57 mm	300 MHz

TABLE IV
COMPARISON OF CUSTOM AND COMPILED FILTERS

Filter Structure	Compiler	Original Core Area	GenFIR Core Area	GenFIR Area / Original Area
QAM Modulator	Hand Layout	5.88 mm × 2.93 mm	6.76 mm × 2.41 mm	94.6%
QAM Demodulator	Hand Layout	5.43 mm × 3.42 mm	4.83 mm × 3.83 mm	99.6%
X/SIN X Predistortion Filter	FIRGEN	3.27 mm × 1.35 mm	1.88 mm × 1.58 mm	67.3%
Half-Band Filter	LPFIR	5.25 mm × 4.50 mm	5.38 mm × 2.64 mm	60.1%

the critical path delay is four adders, and the expected maximum clock frequency of the subfilters is 100 MHz.

VII. CONCLUSION

Architectures for reduced complexity FIR filters were examined in depth, and it was found that a transpose direct form structure with CSD multipliers provides a reasonable trade-off between speed and complexity while retaining the regularity which is desirable for implementation by compiler. With the use of carry-save addition and simple pipelining, this structure can achieve a 2-bit adder delay critical path. With such a short critical path, second-order effects, such as loading of the data broadcast, become a significant portion of the delay, so techniques were presented to reduce this loading to a minimum by both the elimination of the sign extension problem and the automated removal of circuitry which is not strictly required by the particular coefficient set. A new technique for the exploitation of the mirror symmetric properties of subfilter pairs in polyphase filter banks derived from linear phase prototype filters was described. This technique demonstrates how the full advantages of the coefficient symmetry in the prototype filter can be ex-

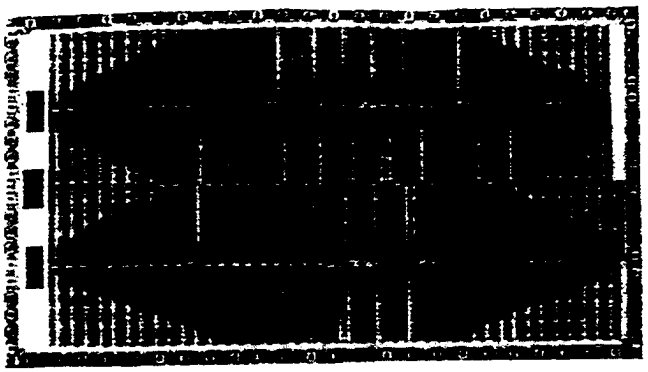


Fig. 10. 320-tap interpolate-by-8 filter.

ploited in the context of the polyphase structure. Finally, the evolution of the filter compiler GenFIR which utilizes all of these techniques was described, and comparisons were made between GenFIR generated filters and similar handcrafted and compiled filters which have been previously published. It was found that GenFIR achieves nearly identical results to the handcrafted layouts and shows significant advances over earlier compilers. An example filter was compiled which places 320 16-bit taps on 0.8 cm^2 of die area. The filter core can operate at sample rates in excess of 100 MHz. Therefore, this example filter has an effective computational power of 32 billion multiply-accumulate operations per second, or about three orders of magnitude larger than can be achieved by a general purpose programmable DSP such as the TMS320C50 (28.6 MOPS) [27], which would occupy a similar or larger die area, and more than two orders of magnitude larger than a commercial programmable filter such as the Harris HSP43881 (240 MOPS) [1], which offers only 8-bit wordlengths. Thus, in applications where a dedicated architecture for FIR filtering can be used, a several-hundred to thousand-fold increase in performance can be achieved with respect to the use of programmable DSP processors.

Acknowledgment

The authors would like to thank Broadcom, Corp. for providing cell libraries, technology files, and design support for the GDT tools.

References

[1] *Harris Digital Filters*, Harris Semiconductor Corp., BR-DSP-001, 1990.

[2] M. Hatamian and S. K. Rao, "A 100 MHz 40-tap programmable FIR filter chip," in *Proc. IEEE Int. Symp. Circuits Syst.*, 1990, pp. 3053–3056.

[3] K. Y. Khoo, A Kwentus, and A. N. Willson, Jr., "An efficient 175 MHz programmable FIR digital filter," in *Proc. IEEE Int. Symp. Circuits Syst.*, 1993, vol. 1, pp. 72–5.

[4] W. Haberecht, E. DeMan, and M. Schulz, "A programmable 32 tap digital interpolation filter in 1.5 μm CMOS with 80 MHz output data rate," in *Proc. IEEE Cust. Int. Circuits Conf.*, 1990, pp. 13.1/1–4.

[5] C. Young and D. L. Jones, "Area-efficient VLSI implementation of FIR digital filters using shifted partial products," in *Proc. 25th Asilomar Conf. on Signals, Systems and Computers*, 1992, vol. 1, pp. 398–402.

[6] G. Casagrande, A. Chiari, C. Golla, and S. Miceli, "VLSI programmable digital filter for video signal processing," *J. VLSI Signal Processing*, pp. 219–231, Dec. 1993.

[7] L. R. Tate, "Sparse canonical signed digit FIR filter," in *Proc. IEEE Cust. Int. Circuits Conf.*, 1987, pp. 215–218.

[8] M. A. Soderstrand and B. Chang, "Design of a high-performance FIR digital filter on a CMOS semi-custom VLSI chip," *Int. J. Mini- and Microcomputers*, vol. 13, pp. 25–30, 1991.

[9] G. Wacey and D. R. Bull, "POFGEN: a design automation system for VLSI digital filters with invariant transfer function," in *Proc. IEEE Int. Symp. Circuits Syst.*, 1993, pp. 631–634.

[10] E. Bidet, C. Joanblanq, and P. Senn, "GENRIF: An integrated VLSI filter compiler," in *Proc. Eur. Conf. on Design Automation*, 1993, pp. 466–471.

[11] M. Ishikawa *et al.*, "Automatic layout synthesis for FIR filters using a silicon compiler," in *Proc. IEEE Int. Symp. Circuits Syst.*, 1990, pp. 2588–2591.

[12] R. Crochiere and L. Rabiner, *Multirate Digital Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1983.

[13] P. Reutz, "The architectures and design of a 20-MHz real-time DSP chip set," *IEEE J. Solid-State Circuits*, vol. 24, pp. 338–348, Apr. 1989.

[14] W. Ulbrich and T. Noll, "Design of dedicated MOS digital filters for high-speed applications," in *Proc. IEEE Int. Symp. Circuits Syst.*, 1985, pp. 255–258.

[15] T. G. Noll and S. Meier, "A 40 MHz programmable semi-systolic transversal filter," *ISSCC Dig. Tech. Papers*, pp. 180–181, Feb. 1986.

[16] S. K. Tewksbury *et al.*, "FIR digital filters for high sample rate applications," *IEEE Comm. Mag.*, vol. 25, pp. 62–72, July 1987.

[17] K. Hwang, *Computer Arithmetic, Principles, Architecture, and Design*. New York: Wiley, 1979.

[18] H. Samueli, "An improved search algorithm for the design of multiplierless FIR filters with powers-of-two coefficients," *IEEE Trans. Circuits Syst.*, vol. 36, pp. 1044–1047, July 1989.

[19] J. Sklansky and M. Lehman, "Ultimate-speed adders," *IRE Trans.*, vol. EC-12, no. 2, pp. 142–148, Apr. 1963.

[20] R. Jain, P. T. Yang, and T. Yoshino, "FIRGEN: a computer-added design system for high performance FIR filter integrated circuits," *IEEE Trans. Signal Processing*, vol. 39, no. 7, pp. 1655–1667, 1991.

[21] J. Laskowski and H. Samueli, "A 150-MHz 43-tap half-band FIR digital filter in 1.2-μm CMOS generated by silicon compiler," in *Proc. Custom Integrated Circuits Conference*, May 1992, pp. 11.4.1–11.4.4.

[22] *Lager Tool Set Manual*. Mississippi State University, 1991.

[23] *L Database and Language User's Guide*. Mentor Graphics Corporation, San Jose, CA, 1990.

[24] T. Lin and H. Samueli, "A 200-MHz CMOS $x/\sin(x)$ digital filter for compensating D/A converter frequency response distortion," *IEEE J. Solid-State Circuits*, vol. 26, pp. 1278–1285, Sept. 1991.

[25] B. C. Wong and H. Samueli, "A 200-MHz all-digital QAM modulator and demodulator in 1.2-μm CMOS for digital radio applications," *IEEE J. Solid-State Circuits*, vol. 26, pp. 1970–1979, Dec. 1991.

[26] R. Hawley, T. Lin, and H. Samueli, "A 200-MHz double-sideband to single-sideband converter in 1-μm CMOS generated by silicon compiler," in *Proc. VLSI Circuits Symp.*, June 1992, pp. 72–73.

[27] *TMS320C5x User's Guide*, Texas Instruments Inc., 2547301-9721 revision B, Sept. 1990.

Robert Hawley (S'91) was born in Flint, MI, on October 26, 1966. He received the B.S. degree from the California Institute of Technology, Pasadena, in 1989 and the M.S. degree from the University of California, Los Angeles (UCLA), in 1991, both in electrical engineering. He is currently nearing completion of the Ph.D. degree in electrical engineering also at UCLA.

In 1993, he joined Broadcom Corporation, Irvine, CA, where he is currently a Manager for Broadband Communications VLSI, working in the area of QAM modem design. His research interests include digital signal processing, digital filter design, adaptive equalization, VLSI architectures for communications systems, and CAD techniques for the implementation of high performance DSP functions.

Bennett C. Wong (S'85-M'93) was born in Los Angeles, CA, on February 2, 1963. He received the B.S. (*summa cum laude*), M.S., Engr., and Ph.D. degrees in electrical engineering from the University of California, Los Angeles, in 1984, 1985, 1988, and 1993, respectively.

From 1985 to 1991 he was with TRW, Inc., Redondo Beach, CA, where he was a Member of the Technical Staff in the Digital Development Laboratory. In 1991, he joined PairGain Technologies, Inc., Tustin, CA, where he is an Engineering Project Manager. He is currently involved with the development of high bit-rate digital subscriber line (HDSL) and asymmetric digital subscriber line (ADSL) technologies. His research interests are in the areas of telecommunications networks, VLSI for digital signal processing, computer arithmetic, signal processing for telecommunications, and high-speed CMOS integrated circuit design.

Thu-Ji Lin (S'88-M'92) received the B.S. degree in electrical engineering from the National Taiwan University in 1984. He received the M.S. and Ph.D. degrees in electrical engineering from the University of California, Los Angeles, in 1989 and 1992, respectively.

Between 1984 and 1986, he served as an Electronics Instructor at the Chinese Army Officer Academic School in Taiwan. From 1986 to 1992, as a Research Associate in the Department of Electrical Engineering at UCLA, he worked on the architecture of digital tunable modulator/demodulator and the implementation of several high-speed CMOS digital filter IC's for communication transmission applications. He joined PairGain Technologies, Inc., Tustin, CA, in 1993. Since then he has been engaged in the development of various digital subscriber line transmission technologies, including HDSL, ADSL, and VDSL. His research interests are in the areas of developing wire and wireless communication systems and applying DSP algorithms and VLSI technologies to their implementation.

Joe L. Laskowski received the B.S. degree in computer science from Texas University in 1988. He received the M.S. degree in electrical engineering from University of California, Los Angeles, in 1991, concentrating in the area of integrated circuits and systems.

From 1991 to 1995, he worked for Lattice Semiconductor, Milpitas, CA, designing programmable logic devices. He is presently at Broadcom Corporation, Irvine, CA where he is working on high-speed integrated circuits for communication systems. His interests include high-speed integrated circuits and systems.

Henry Samueli (S'75-M'81) was born in Buffalo, NY, on September 20, 1954. He received the B.S., M.S., and Ph.D. degrees in electrical engineering from the University of California, Los Angeles (UCLA), in 1975, 1976, and 1980, respectively.

From 1980 to 1985, he was with TRW, Inc., Redondo Beach, CA, where he was a Section Manager in the Digital Processing Laboratory of the Electronics and Technology Division. His group was involved in the hardware design and development of military satellite and digital radio communication systems. From 1980 to 1985 he was also a part-time instructor in the Electrical Engineering Department at UCLA. In 1985 he joined UCLA full-time where he is currently a Professor in the Electrical Engineering Department. His research interests are in the areas of digital signal processing, communications systems engineering, and CMOS integrated circuit design for applications in high-speed data transmission systems. In 1988, he co-founded PairGain Technologies, Tustin, CA, a telecommunications equipment manufacturer, and in 1991, he co-founded Broadcom Corporation, Irvine, CA, a communications IC supplier.

Dr. Samueli is the recipient of the 1988/1989 TRW Excellence in Teaching Award of the UCLA School of Engineering and Applied Science and the Meritorious Paper Award of the 1991 Government Microcircuit Applications Conference.

A 900 MHz CMOS Frequency-Hopped Spread-Spectrum RF Transmitter IC

A. Rofougaran, G. Chang, J.J. Rael, M. Rofougaran, S. Khorram, M.-K. Ku, E. Roth, A.A. Abidi, and H. Samueli

Proceedings of the IEEE Custom Integrated Circuits Conference
pp. 209-213
May 1996

A 900 MHz CMOS Frequency-Hopped Spread-Spectrum RF Transmitter IC

A. Rofougaran, G. Chang, J. J. Rael, M. Rofougaran, S. Khorram, M-K. Ku, E. Roth, A. A. Abidi, and H. Samueli

Electrical Engineering Department
University of California
Los Angeles, CA 90095-1594

Introduction

This paper reports on the first monolithic CMOS RF transmitter, which is part of a complete low-power transceiver operating in the 902-928 MHz ISM frequency band [1]. The system can transmit data at rates up to 160 kb/s by frequency-shift keying a carrier which is periodically hopped according to a pseudorandom code across 26 MHz. The key elements to accomplish this in the transmitter are: an *agile* baseband *frequency synthesizer*, a low-noise 915 MHz *local oscillator*, *upconversion mixers*, some needed *RF filters*, and a *power amplifier*. All these elements are integrated on a 1-μm CMOS IC (Fig.1(a)).

Transmitter Architecture

At the core of the transmitter architecture is the agile frequency synthesizer. Hopping frequency synthesizers are conventionally based on a phase-locked loop (PLL) with a digitally-controlled variable divider. Finely separated frequencies may be synthesized from a fixed crystal-based reference by operating the divider in fractional-N mode. This approach suffers from two limitations: the finite PLL unity-gain bandwidth, often about one-tenth of the reference frequency, results in a long loop settling-time when the divider modulus is switched; and it is difficult to make a VCO with a wide tuning range, whose output spectrum is also acceptably free of harmonics and spurious frequencies over that range. We have instead adopted a solution based on direct-digital frequency synthesis (DDFS) to overcome both these problems. A 12-bit input frequency control word sets the accumulation rate in a register, which addresses ROMs containing coefficients of the trigonometric sine and cosine waveforms (Fig.2) to produce digital-domain quadrature outputs at the programmed frequency. The accuracy of ROM coefficients determines the harmonic and spurious frequency content, guaranteed to be -72 dBc or less in this implementation. The DDFS output is followed by a 10-bit D/A converter (DAC) to obtain an acceptably low harmonic distortion and noise at the synthesizer output. We have previously reported measurements on a prototype of this agile frequency synthesizer [2]. Frequency-shift keyed modulation is superimposed on the hopping carrier by simply adding or subtracting a small, fixed offset to the frequency-control word.

The quadrature outputs of the DDFS/DAC are upconverted by quadrature phases of a fixed frequency 915 MHz local oscillator (LO), and added or subtracted to select either the upper or the lower side-frequency. In this way, any frequency may be synthesized in the 902-928 MHz band, to within the resolution of the 12-bit frequency control word. DAC and subsequent nonlinearities will set the important spurious frequencies in the output spectrum. This arrangement requires a *fixed-frequency* 915 MHz LO, which may be entirely integrated and then embedded in a wideband PLL to reduce phase noise.

Each upconversion mixer in the quadrature channels comprises a four-FET commutating switch. With a sufficiently large LO voltage applied at the FET gates, this mixer upconverts a complex baseband spectrum with little intermodulation distortion. An on-chip bandpass filter suppresses signal energy upconverted by the 3rd LO harmonic.

The local oscillator consists of a pair of cross-coupled *LC* oscillators, which synchronize into quadrature phases [3] (Fig.3). A source-follower FET with a voltage-controlled PMOS resistor in series with its gate synthesizes a load which appears inductive at high frequencies. This resistor tunes the LO frequency. The LO output phase noise spectral density is -85 dBc/Hz at a 100 kHz offset from the carrier.

The UCLA Low-Power Wireless Transceiver Project is supported by the US Advanced Research Projects Agency, and by a consortium of semiconductor companies under the State of California MICRO Program. For more information, e-mail abidi@icsl.ucla.edu

A prototype of the on-chip power amplifier has been previously published [4]. Under control of a digital word, it delivers power in the range of –15 dBm (30 μW) to +15 dBm (30 mW) to an antenna load in increments of 1 dB. The power amplifier comprises an inductively loaded common-source FET preamplifier, which drives a binary-weighted array of FETs whose drains tie together into the off-chip load (Fig.4). The array of FETs is biased close to threshold, and attains a peak power-conversion efficiency of about 40% when amplifying a single-tone input. Two halves of the amplifier drive the antenna in balanced mode.

Circuit Design

The frequency spectrum of the discrete-time DDFS analog output contains images of the main tone around the clock frequency. As shown in Fig.1(b), the single-sideband upconversion mixer selects the images ❷ and ❸ separated from the main tone ❶ by exactly the clock frequency, while suppressing the other image in much the same way that it cancels the unwanted sideband of the main tone. These images must not radiate from the antenna. They may be removed after the DAC with a lowpass filter of sufficiently high order, but with a 20 MHz cutoff frequency, it is very difficult to build an active filter with a 60 dB linearity commensurate with the DAC output.

Instead, the DDFS/DAC clock frequency is raised to 80 MHz, and the upconverted tone and the image frequencies are applied to the power amplifier with no prior filtering. All harmonics and intermodulation tones that nonlinearity in the power amplifier creates are also separated at least by 80 MHz, and fall well into the stopband of the three-resonator, passive, off-chip dielectric resonator filter connected to the amplifier output (Fig.1(b)). With a stopband attenuation of 40 dB or greater, this passive filter rejects the images without distorting the signal. Owing to the $\sin x/x$ rolloff in the spectrum of the sampled waveform, the two main image tones, ❷ and ❸, are at least 15 dB lower than the main tone, and do not overload the power amplifier to the point that it suffers a loss in efficiency.

The quasi-passive DAC core consists only of capacitors and switches (Fig.5(a)), and readily operates with a three-phase clock at 80 MHz. In the earlier prototype the op amp charge-to-voltage buffer required two additional clock phases, and its settling time became the main bottleneck to higher clock rates. We have reconfigured it to re-use two of the DAC clock phases, and now this too clocks at 80 MHz.

The DAC, through the four-FET upconversion mixer, must drive the input capacitance of the on-chip power amplifier with a 0.5V full-scale signal. This capacitance is too large a load for the op amp, so a second open-loop buffer consisting of a differential pair degenerated by a polysilicon resistor is used (Fig.5(b)). The spurious-free dynamic range at the DDFS/DAC output, prior to the second buffer, is about 60 dB [2]. Although the buffer drains a relatively small current, it adds –45 dB 3rd harmonic distortion to the full-scale DDFS output, and is the weak point in overall system linearity.

The two four-FET switch mixers are shorted together at the outputs (Fig.6) to select one sideband. The sidebands upconverted by the 3rd harmonic of the LO must be adequately suppressed, otherwise they will alias as spurious tones within the ISM band after suffering power amplifier nonlinearity. A low-Q bandpass filter, consisting of an on-chip 50 nH spiral inductor in series the input capacitance of the power amplifier, adequately suppresses signals at 2.7 GHz. The spiral inductor is suspended above a cavity in the substrate (Fig.7) [5], which greatly reduces the parasitic capacitance and increases its self-resonance to beyond 2.7 GHz.

All the above circuits operate on a 3V power supply.

Experimental Results

The entire transmitter is fabricated in a 1-μm CMOS, double-metal technology with a linear capacitor. It occupies an active area of 6×3.8 mm (Fig.8) mainly consumed by the large on-chip inductors. The circuit, except the LO, drains 55 mA when operating at speed. The LO with spiral inductor loads drains 10 mA, while in this version with simulated inductors the current drain is 30 mA.

The transmitter output spectra are evaluated at a mid-scale power amplifier setting of about +5 dBm (3 mW). A 10 MHz DDFS/DAC tone shows several harmonics in the upconverted spectrum at about –50 dBc, which is also the LO leakage level (Fig.9). The unwanted sideband is roughly at –43 dBc. This performance is either at par or superior to sideband selection in previously reported integrated high-f_T bipolar transmitter ICs for GSM [6-8], and implies an on-chip quadrature inaccuracy of less than 1°.

The transmitter is also evaluated by hopping the DDFS in the lower-half ISM band, and measuring the average levels of the unwanted sidebands lying in the upper-half band, which will interfere with other users in that part of the band (Fig.10). All the unwanted sidetones are 46 dB lower than the wanted signal, except one which is only 40 dB lower. This is an acceptable sideband

suppression in a microcell-based system where transmitted power is controlled over a 30 dB range. No off-chip filters are inserted at the power amplifier output for these measurements.

These results show that with the appropriate mix of architecture and circuits, it is possible to attain a a level of integration and performance in a modest 1-μm CMOS technology comparable to more advanced RF silicon technologies, with no apparent untoward effects of parasitic coupling through the substrate or other observable detriments.

References

1 J. Min, A. Rofougaran, H. Samueli, and A. A. Abidi, "An All-CMOS Architecture for a Low-Power Frequency-Hopped 900 MHz Spread-Spectrum Transceiver," presented at *Custom IC Conf.*, San Diego, CA, pp. 379-382, 1994.

2 G. Chang, A. Rofougaran, M. K. Ku, A. A. Abidi, and H. Samueli, "A Low-Power CMOS Digitally Synthesized 0-13 MHz Agile Sinewave Generator," presented at *Int'l Solid State Circuits Conf.*, San Francisco, pp. 32-33, 1994.

3 A. Rofougaran, J. Rael, M. Rofougaran, and A. A. Abidi, "A 900 MHz CMOS LC Oscillator with Quadrature Outputs," to be presented at *Int'l Solid State Circuits Conf.*, San Francisco, CA, 1996.

4 M. Rofougaran, A. Rofougaran, C. Olgaard, and A. A. Abidi, "A 900 MHz CMOS RF Power Amplifier with Programmable Output," presented at *Symp. on VLSI Circuits*, Honolulu, pp. 133-134, 1994.

5 J. Y.-C. Chang, A. A. Abidi, and M. Gaitan, "Large Suspended Inductors on Silicon and their use in a 2-μm CMOS RF Amplifier," *IEEE Electron Device Letters*, vol. 14, no. 5, pp. 246-248, 1993.

6 W. Veit, J. Fenk, S. Ganser, K. Hadjizada, S. Heinen, H. Herrmann, and P. Sehrig, "A 2.7V 800MHz-2.1GHz Transceiver Chipset for Mobile Radio Applications in 25 GHz f_T Si-Bipolar," presented at *Bipolar Circuits & Technology Mtg.*, Minneapolis, pp. 175-178, 1994.

7 T. Stetzler, I. Post, J. Havens, and M. Koyama, "A 2.7V to 4.5V Single-Chip GSM Transceiver RF Integrated Circuit," presented at *Int'l Solid State Circuits Conf.*, San Francisco, pp. 150-151, 1995.

8 C. Marshall, F. Behbahani, W. Birth, A. Fotowat, T. Fuchs, R. Gaethke, E. Heimerl, S. Lee, P. Moore, S. Navid, and E. Saur, "A 2.7V GSM Transceiver IC with On-Chip Filtering," presented at *Int'l Solid State Circuits Conf.*, San Francisco, pp. 148-149, 1995.

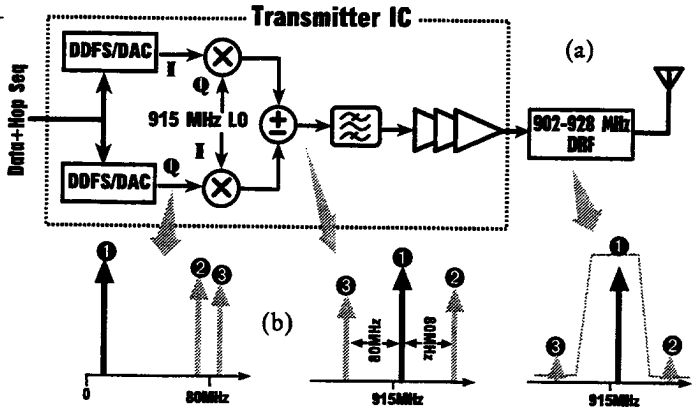


Fig.1 (a) Block diagram of transmitter IC. (b) Frequency spectrum of a single tone generated by DDFS in various parts of the transmitter. The output dielectric resonator filter acts as an image-suppression mixer.

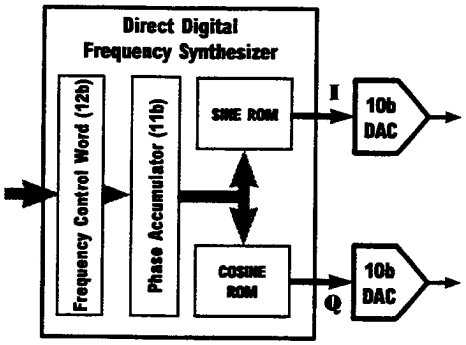


Fig.2 Block diagram of DDFS/DAC unit

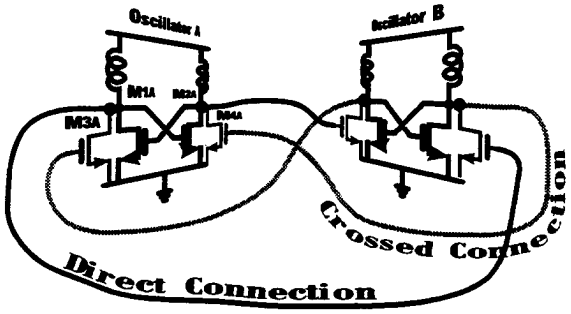


Fig.3 LC oscillator which produces balanced outputs in quadrature.

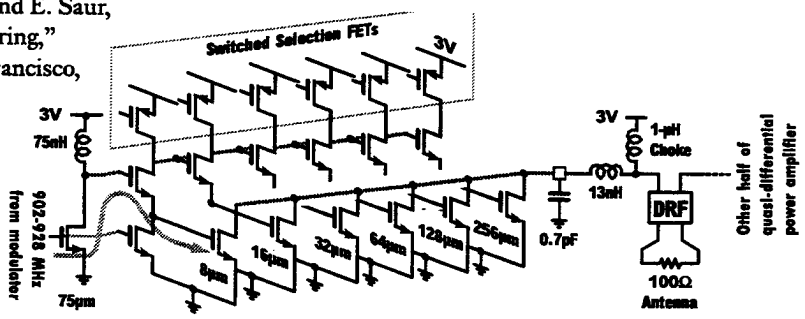


Fig.4 Power amplifier consists of binary-weighted array of FETs.

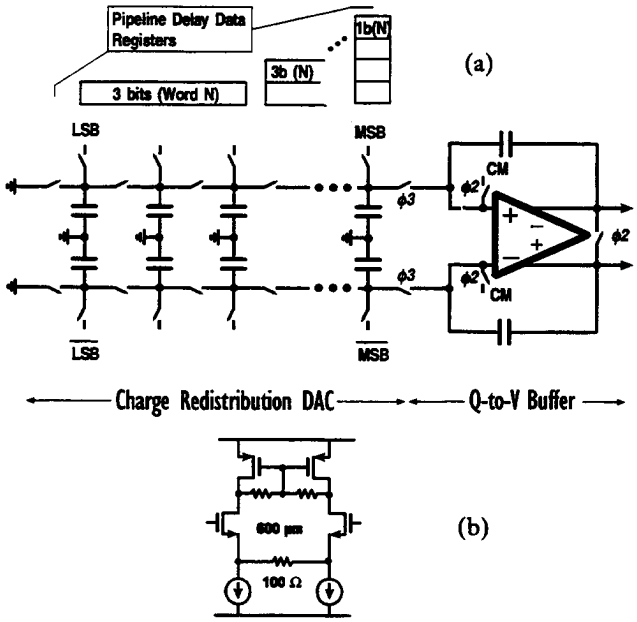


Fig.5 (a) Quasi-passive 10-bit pipelined charge redistribution DAC; (b) Additional buffer to drive power amplifier through upconversion mixers.

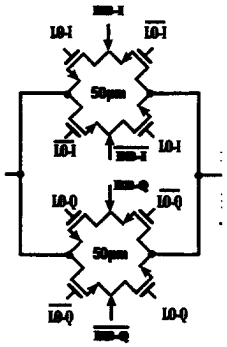


Fig.6 Four-FET switch upconversion mixers in I and Q channels sum outputs as currents.

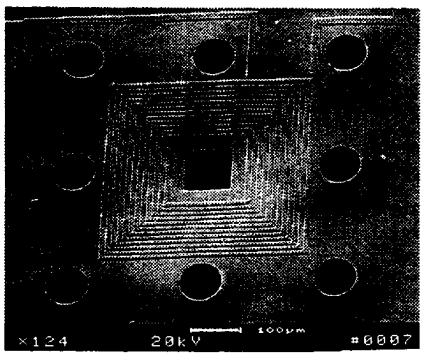


Fig.7 A large-inductance spiral suspended above a cavity on a CMOS substrate. The etchant isotropically excavates a cavity under the exposed holes surrounding the spiral.

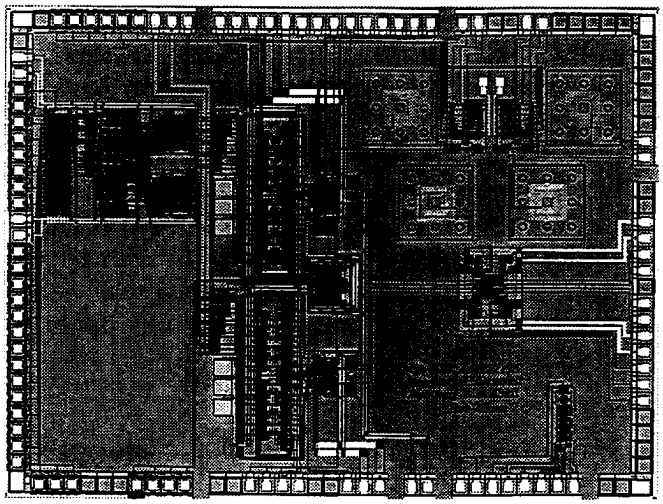


Fig.8. Transmitter IC microphotograph

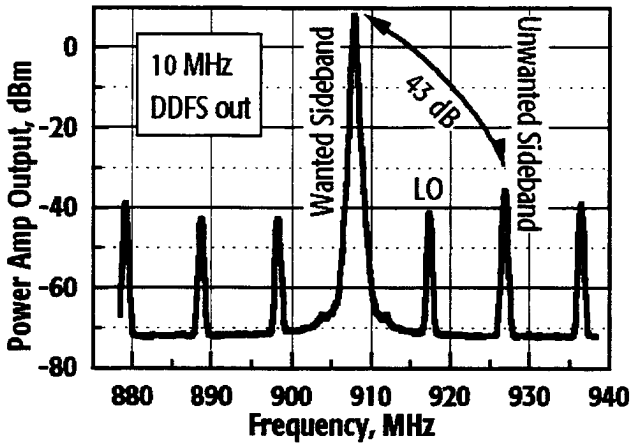


Fig.9 Measured output spectrum from power amplifier, with single 10 MHz tone generated by DDFS.

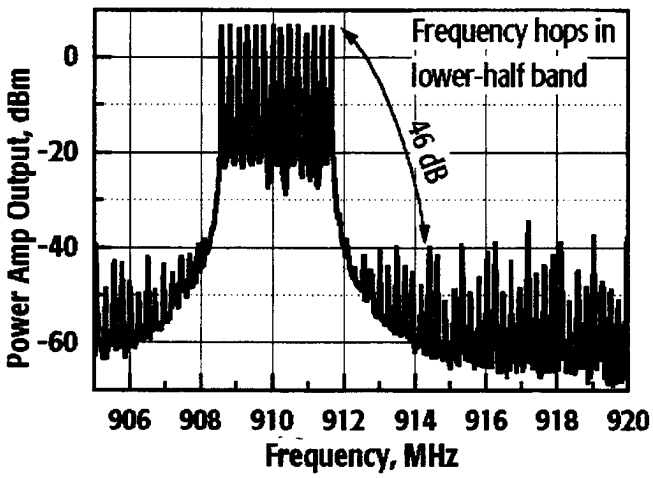


Fig. 10 Measured output spectrum from power amplifier, with DDFS hopping over 14 frequency slots in lower half of 902-928 MHz band.

A Monolithic 900 MHz Spread-Spectrum Wireless Transceiver in 1-μm CMOS

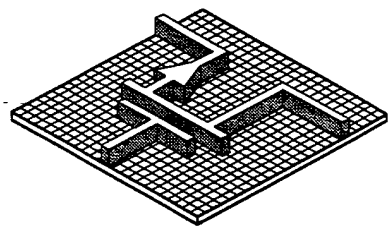
A.A. Abidi, A. Rofougaran, G. Chang, J. Rael, J. Chang, M. Rofougaran, P.J. Chang, and S. Khorram

presented at the Workshop on Advances in Analog Circuit Design
April 1996

UCLA

A Monolithic 900 MHz Spread-Spectrum Wireless Transceiver in 1-µm CMOS

Presented at the *Workshop on Advances in Analog Circuit Design* (AACD), Lausanne, Switzerland, April 2-4, 1996
(Proceedings are limited in circulation to the workshop participants)



A Monolithic 900 MHz Spread-Spectrum Wireless Transceiver in 1-µm CMOS*

Asad A. Abidi

with

Ahmadreza Rofougaran Glenn Chang

Jacob Rael James Chang

Maryam Rofougaran Paul Jinyun Chang

Shahla Khorram

Electrical Engineering Department

University of California

Los Angeles, CA 90095-1594 (USA)

abidi@icsl.ucla.edu http://www.icsl.ucla.edu/aagroup

TABLE OF CONTENTS

Table of Contents ____ *1*

Introduction ____ *3*

System Architecture ____ *4*

Detailed Circuit Description ____ *7*

Digital Frequency Synthesizer ____ **8**

Upconversion Mixers and Single-Sideband Selection ____ **10**

RF Power Amplifier ____ **12**

Local Oscillator with Quadrature Outputs ____ **14**

Integrated Transmitter ____ **17**

Technology for Large-Value On-Chip Inductors ____ **17**

Low-Noise Amplifier and Downconversion Mixer ____ **18**

Channel-Select Lowpass Filter ____ **20**

Limiting Amplifier and Detector ____ **22**

* The UCLA Low-Power Transceiver Project is financially supported by the US Advanced Research Projects Agency, Rockwell Semiconductor Systems, Texas Instruments, Harris Semiconductor, Advanced Micro Devices, and the State of California MICRO Program, with equipment contributions from Hewlett-Packard Company.

References _____ *24*

INTRODUCTION

The RF and IF sections of wireless communications devices have traditionally comprised a collection of discrete active and passive components, while IC technology has made an impact on the baseband sections. The needs for low power operation and greater miniaturization impel RF and IF circuit technology towards greater levels of integration [1]. The single-chip radio has yet to be realized in the 1 to 2 GHz frequency band, where most of the digital cellular and other widespread data communications take place today. Such a chip would connect to the antenna on one end, and on the other end provide ports for the input and output of baseband data. At the current level of interest in wireless ICs, though, such an integrated radio is expected soon.

There are some historical hurdles to be overcome on the way to realizing the single-chip radio, others which are technological. Until recently, the small group of RF circuit practitioners was mainly trained in the discrete art, and they would normally implement circuits to the specifications of another group of practitioners, the system or radio architects. There was usually little exchange between the two groups, not least because they may have lacked a common language of communication. Miniaturization of the radio was usually a result of packaging the discrete components into smaller form-factors, as was demonstrated by the early generations of the Sony Walkman. The one significant break from this RF tradition appeared in the development of the integrated radio paging receiver in the 1970s, where it may be said that the style of baseband analog IC design was, in its full sense, first applied to a radio-frequency device [2-4]. The technological hurdle arises from the mostly perceived, although sometimes real, inability of well-established, high-volume IC technologies capable of mixed analog-digital integration, namely silicon CMOS, bipolar, and BiCMOS, to implement RF functions.

This is a progress report on one of the first coordinated, large-scale research efforts towards realizing a single-chip 900 MHz digital radio. Several factors are responsible for what has been achieved so far. First, this was an example of a *simul-*

taneous evolution of the systems architecture with the enabling circuits components, thus allowing for a joint optimization. This luxury is seldom afforded to circuit designers working to the established specifications of an industry standard, but because our transceiver operates in one of the three Industrial, Scientific, and Medical (ISM) frequency bands opened up by the Federal Communications Commission for unlicensed spread-spectrum use, we could design a unique system architecture. A user of these bands is required to spread the transmitted spectrum by a minimum amount, and to transmit no more than a specified power from the antenna, but is otherwise free to use any modulation scheme or spread-spectrum strategy. Second, it was decided at the outset that the *entire transceiver* would be integrated in *CMOS*, which affords an unprecedented flexibility in choosing between an analog or digital implementation of the various receiver blocks. Third, and its importance cannot be underemphasized, the project evolved from a close collaboration between analog and digital circuit designers, communication system engineers, and antenna designers.

SYSTEM ARCHITECTURE

The key transceiver specifications are listed in Table 1, and the rationale for their choice are then described below [5].

Table 1: Transceiver Specifications

Operating Band	902-928 MHz	*Data Rate*	Up to 160 kb/s
Modulation Scheme	Binary or Quaternary FSK	*Spreading Scheme*	Frequency-Hop

The data rate is sufficiently high to accommodate voice and ISDN-type data, without requiring power-hungry equalization of typical delay-spreads in the receiver. A frequency-hopped spectrum spreading method is used over the more well-known direct-sequence, because it is possible with this method to cover an arbitrarily wide frequency range, without the attendant increase in rate of trans-

mitted symbols, or chips. This results in a lower power receiver front-end. Frequency-hopping also enables the use of FSK modulation, which in turn makes it possible to use a direct-conversion receiver, whose virtues of high integration and low power are by now well-known [6].

A block diagram of the transceiver is shown in Figure 1. This shows a single transmitter, and two-branch diversity with two receive channels whose outputs

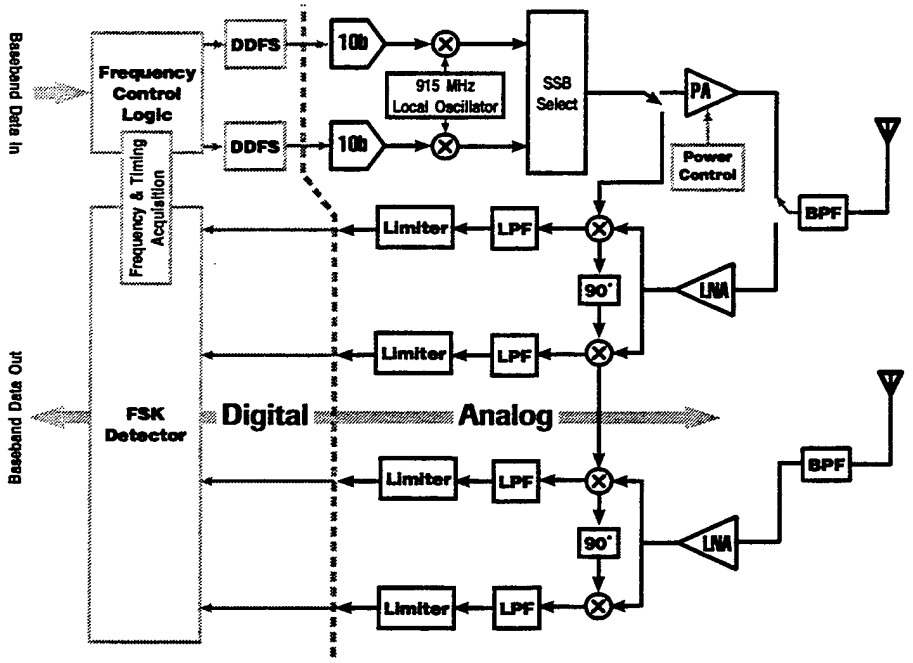


Figure 1: Transceiver Architecture

combine in the detector. As is characteristic of communications by spread-spectrum, transmission and reception take place in the same frequency band, and therefore one RF preselect bandpass is used bidirectionally. This dielectric resonator filter with a passband from 902 to 928 MHz is the only discrete component shown on this block diagram; not shown is the only other precision discrete component, a crystal with a resonant frequency at about 10 MHz which fixes the 915 MHz frequency of the local oscillator.

At the core of the transceiver is an *agile frequency source* based on direct-digital frequency synthesis (DDFS). This is a table lookup technique, whereby an accu-

mulator which ramps at a rate set by an input frequency control word addresses a sinewave ROM to produce a discrete-time digital sinewave at that frequency. The absolute accuracy of the output frequency is then set by the fixed DDFS clock, its resolution by the input wordlength, and the wordlength of the internal arithmetic as well as the ROM construction determine the spectral purity of the digital sinewave. This approach gives instant frequency agility with continuous phase, but it requires a linear D/A converter, and an analog filter to suppress the images around the clock frequency. The spurious tones in the DDFS output are at least 76 dB below the main tone, so to convert it into a discrete-time analog sinewave with commensurate spectral purity, the DDFS is followed by a 10b DAC. The DDFS also provides an accurate quadrature output, which is converted to the analog domain by a second identical DAC.

This frequency synthesizer operates on the principle of producing a 26 MHz wide spread-spectrum modulated signal at *baseband*, and then upconverting that spectrum as a block with a fixed 915 MHz LO centered at in the middle of the ISM band. Quadrature phases of the LO upconvert the quadrature baseband spectrum into either the upper or lower half of the ISM band, depending on whether the two arms add or subtract. Thus, a DDFS output tone lying in the interval 0 to 13 MHz may be upconverted into either the upper or lower 13 MHz of the ISM band, depending on the sign of the summation, which is itself implemented as a digital phase reversal in one of the two DDFS quadrature channels.

Data modulates the carrier with a binary or quaternary *frequency-shift keyed* (FSK) *modulation*. The DDFS also serves as the data modulator, because a data 1 or 0 imposes a small, fixed offset corresponding to ±160 kHz (as well as ±80 kHz when quaternary FSK is used) on the frequency control word. The output spectrum is therefore constant-envelope, and amenable to *nonlinear power amplification* for a high conversion efficiency. Harmonics created by such a nonlinear element in the large-signal path lie out of the ISM band, but intermodulation distortion between multiple closely-spaced tones may create spurious in-band products. A fre-

quency-hopped system with FSK modulation produces, in principle, only one frequency at a given time, but other effects such as upconversion by harmonics of the LO may cause spurious in-band products. A *passive, on-chip polyphase filter* after the upconversion mixer suppresses these undesirable harmonics, and also improves the extent to which the unwanted sideband is suppressed.

Owing to the choice of FSK modulation, it is feasible to use a *direct-conversion*, or zero-IF, *receiver* [6]. The preselect filter passes a received signal at the antenna comprising all users of the ISM band in the cell, which then passes through the on-chip RF *low-noise amplifier* (LNA). The amplified ISM band is downconverted and de-hopped into two channels by mixers driven by quadrature phases of a frequency-hopped local oscillation, which is synchronized to the hopping pattern of the sought user. The desired channel is thus centered at DC, while all other users in the ISM band lie farther away, up to +13 MHz or –13 MHz. An on-chip active *lowpass filter* selects the desired channel, and suppresses all adjacent users. It also provides baseband gain. With the desired channel isolated, the modulation is recovered by driving the signal into a *limiting amplifier* which asynchronously quantizes the received waveform to 1-bit. A digital *correlating FSK detector* then determines which of the two (or four) possible frequency offsets is being received, and makes a data decision.

The hopping local oscillation for the downconversion mixer is derived from the transmitter's frequency synthesizer, which is switched away from the power amplifier during reception. A digital *frequency and timing acquisition loop* connected to the detector synchronizes the DDFS hopping sequence to the incoming hopping pattern, and aligns the symbol clock in the detector, as well as correcting small errors between the receiver and transmitter crystal references.

DETAILED CIRCUIT DESCRIPTION

The various sub-sections below describe the 1-μm CMOS circuits in each block of the transceiver, following the same order as in the previous section. References in

all the sub-sections direct the interested reader to other publications devoted to each circuit.

Digital Frequency Synthesizer

A block diagram of the direct-digital frequency synthesizer is given in Figure 2. The DDFS frequency control word provides a frequency resolution of 2^{-11} of the

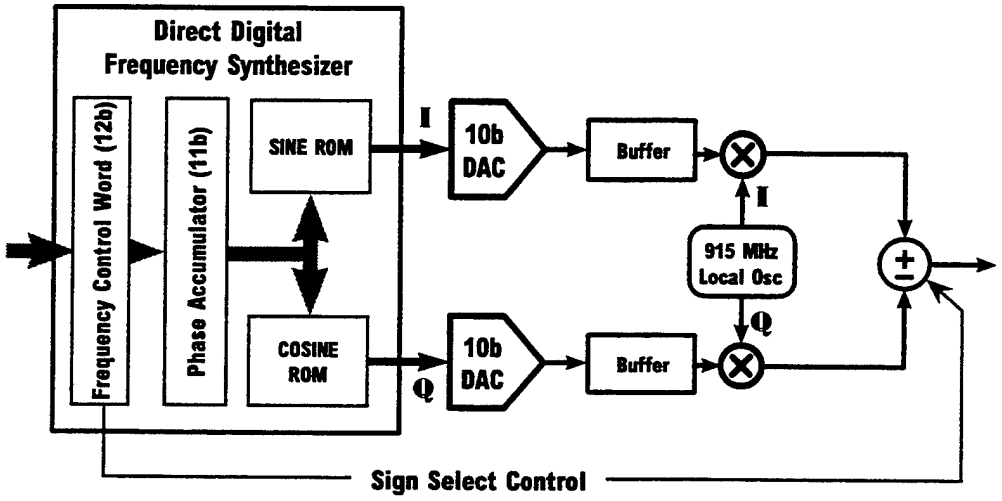
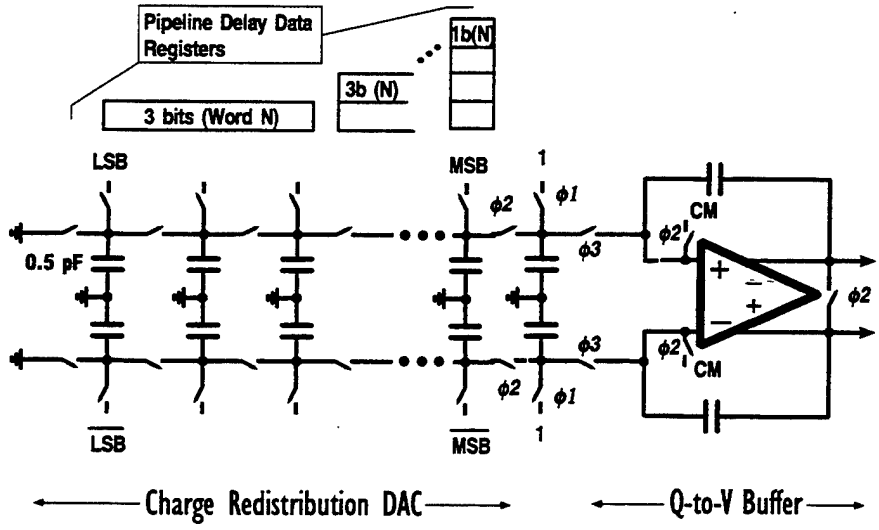


Figure 2: Agile Digital-based Frequency Synthesizer

clock rate. A single quarter-wave ROM stores the difference between amplitude and phase, and comprises coarse and fine tables. Phase-shifted addresses generate both in-phase (I) and quadrature (Q) outputs from the same ROM. These simplifications [7] together yield a 32 times reduction in size compared to a straightforward approach, and this is mainly why a static CMOS implementation dissipates only 60 mW when operating at 80 MHz [8].



A quasi-passive, pipelined charge-redistribution architecture, shown in Figure 3, implements the glitch-free DAC [8]. It consists of ten stages of equal-valued unit capacitors, which operate on a three-phase clock to successively bisect charge according to decreasingly significant bits of a 10-bit input word. The final charge is converted to voltage in a reset op-amp based integrator. The op amp is a gain-boosted cascode differential pair. The DAC core dissipates only 8 mW from 3 V at 80 MHz, principally in the clock drivers. This simple arrangement yields very pure synthesized output tones, with the highest spurious level of –62 dBc at low synthesized frequencies, rising to -57 dBc worst case at high synthesized frequencies (Figure 4) [8]. The output voltage swing is 0.5V ptp differential.

This high sinewave purity is sacrificed in the buffer following the DAC, which

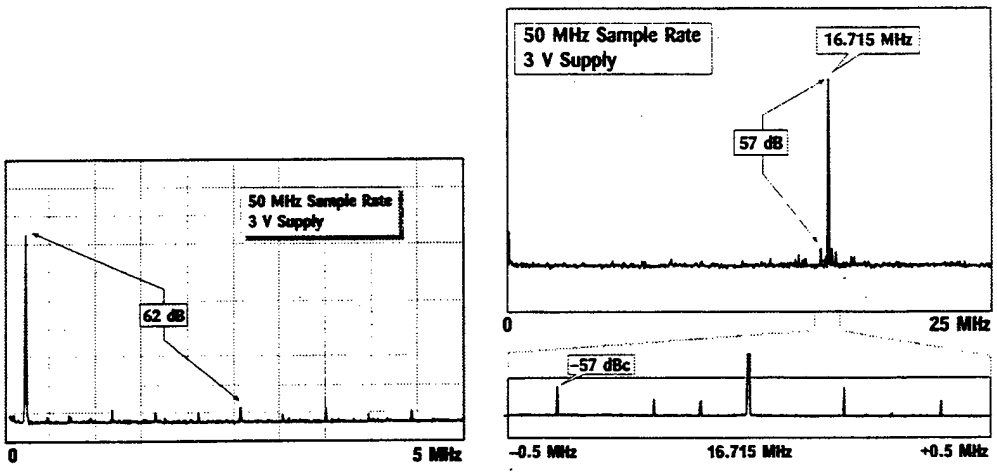


Figure 4: Measured spectral outputs of DDFS/DAC

drives the power amplifier through the series switches of the upconversion mixer. An op amp-based solution to this would be too power-hungry, so a simple differential pair is used, resistor-degenerated at the sources to accept the large DAC output with relatively low distortion. The buffer adds a 3rd harmonic of –45 dB, which will upconvert to within the ISM band for synthesized frequencies of less than 4.33 MHz.

Upconversion Mixers and Single-Sideband Selection

Following frequency synthesis at baseband, two four-FET switch mixers upconvert the large baseband sinewaves in the quadrature channels to 915 MHz (Figure 5). This simple mixer is very well-suited to MOS implementation, and when driven with large LO signals at the gate, it upconverts with high linearity and efficiency. However, full commutation by the switches also upconverts by the 3rd and 5th harmonics of the LO, and these additional frequencies may produce aliases through the nonlinear power amplifier which lie in-band.

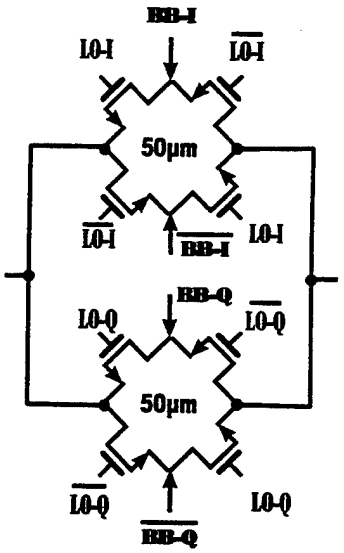


Figure 5: Upconversion FET-switch mixers

Although the mixers are schematically shown here selecting one sideband by summing together the outputs in current-mode, they actually drive separate ports of an RC polyphase filter which further suppresses the unwanted sideband. The polyphase filter (Figure 6) is a generalization of the simple RC-CR filter, and uses phase lag and lead to reinforce one rotational sequence of quadrature phases, say clockwise, while suppressing the other, say counterclockwise [9]. This is a particularly useful addition to quadrature upconverters, because the desired sidebands will assume the opposite rotational sequence to the unwanted sidebands. Furthermore, this is a broadband circuit, and does not require great accuracy in the R and C to be effective. A single polyphase filter section in the transmitter path suppresses the unwanted sideband by another 10 dB after upconversion, and compensates for residual phase errors in quadrature phases of the LO. However, the baseband gains of the two DACs and their buffers in the frequency synthesizer

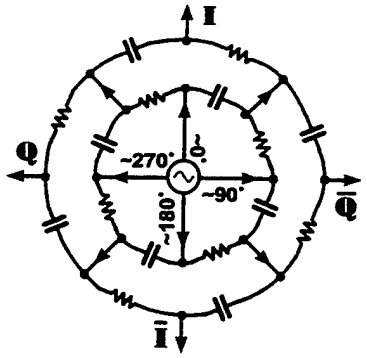


Figure 6: RC Polyphase filter: cascade of two sections

must match to better than 0.1 dB if the unwanted sideband is to be suppressed by more than 45 dB (Figure 7), and the polyphase filter cannot compensate for this source of error.

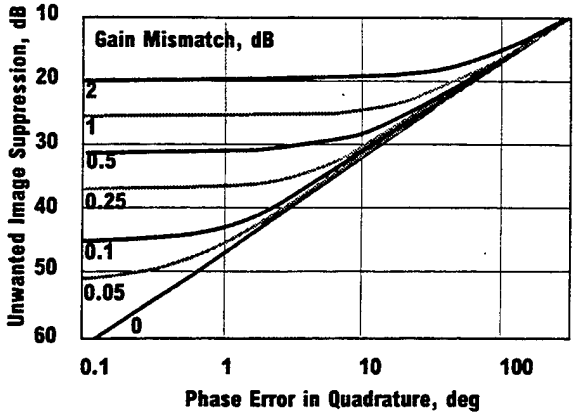


Figure 7: Unwanted sideband suppression with gain and LO phase errors

Owing to its broadband properties, the polyphase filter exerts a fortuitous suppression on the frequencies upconverted by the 3rd harmonic of the LO. When it is driven so as to supplement the inherent selection of the quadrature upconverter, the polyphase filter appears contrary to the 3rd harmonic components, because all the relative angles are now rotated by 3×. Thus, the quadrature upconverter inherently cancels one upconverted sideband at the 3rd harmonic, and this polyphase filter as above cancels the other sideband. The combined effect is to suppress *both* sidebands around the 3rd LO harmonic, which neatly substitutes for an LC notch filter.

It is appropriate here to discuss how the single-sideband upconverter treats the

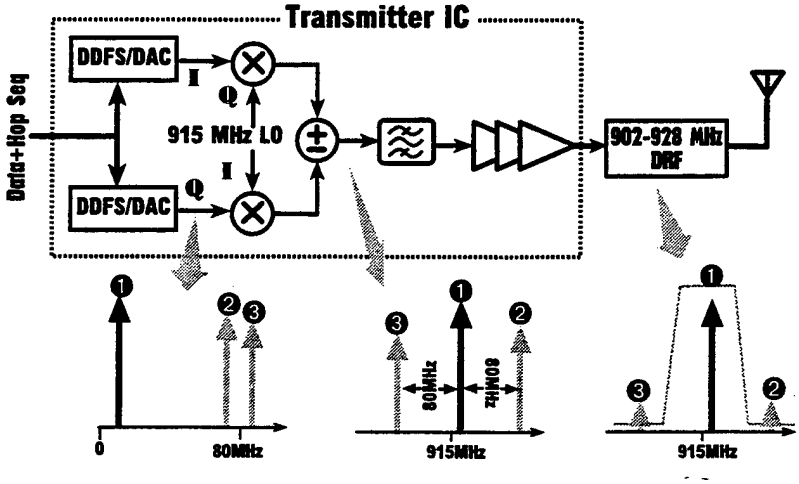


Figure 8: Image frequencies in transmitter

image frequencies around the DDFS/DAC clock frequency. After quadrature up-

conversion, the circuit selects the upper sideband of the lower image, and the lower sideband of the upper image, so that the upconverted images are all separated by exactly the clock frequency (Figure 8), 80 MHz in this implementation. Owing to the *sin x/x* rolloff in the spectrum of the sampled waveform, and the large oversampling between the DAC output of 13 MHz or less with an 80 MHz clock, the images are at least 15 dB lower than the main tone. They are left to propagate through the power amplifier, whose nonlinearity only redistributes energy between the various equally-spaced tones but does not create any new tones. Finally, the off-chip RF preselect filter suppresses the images by an additional 40 dB or so, as the choice of DAC clock frequency forces them to lie in the filter's stopband. Image suppression is therefore accomplished with this passive filter, rather than on-chip filters which would require a prohibitively large power dissipation to handle the large signal levels with sufficiently low distortion.

RF Power Amplifier

Some specifications on the power amplifier in a microcell-based wireless system are relaxed compared to large-cell based systems, while others are more stringent. This power amplifier is integrated on the same 1-μm CMOS substrate as the rest

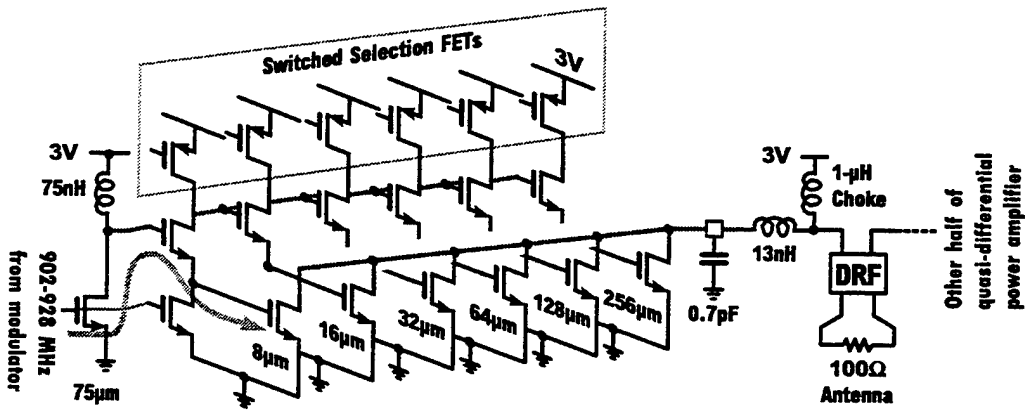


Figure 9: RF Power Amplifier

of the components in this transceiver, as it is required to deliver a maximum of only 30 mW to the output. However, for maximum user capacity, the output

power must also be controllable in 1 dB increments over a range of 30 dB, which corresponds to a minimum of 30 μW.

The power amplifier derives from baseband CMOS circuit design. A binary-weighted array of FETs is connected together at the drains to supply the antenna current (Figure 9) [10]. The FETs are normally biased close to threshold, and are driven into conduction by a large voltage swing at the gates. The requirement for maximum power delivery is met with a 3V supply by providing a push-pull voltage into the antenna with the two halves of the quasi-differential circuit. In turn, a scaled array of source followers drives the gates of the output FETs, and their common gate receives a large swing from a preamplifier with a large on-chip inductor load. This load enables the preamplifier drain voltage to swing up to 5.5 V, thus overcoming the large V_{GS} drop across the source followers. The preamplifier also sets the gate bias of the output FET array through a DC feedback loop.

The output power is controlled by selectively enabling the source followers in the array through PMOS switches in series with the drains. These switches are outside the RF signal path. The effective width of the output FET is thus under direct control of a digital word, while the input voltage to the power amplifier, derived from the upconversion mixer, always stays constant. This reduces output-power dependent pulling of the upconversion LO.

An off-chip inductor matches the capacitive impedance of the power amplifier to 50Ω. This capacitance is dominated by the output bonding pad, and the sum of the drain junction capacitances of the FET array is the next contributor. The latter capacitance does not change much with the power-control word, but the net output resistance contributed by the r_{ds} of the ON FETs in the array does depend on this word. This causes the output reflection coefficient, s_{22}, to vary somewhat but at maximum power it attains a minimum value of –30 dB by design. The amplifier is unconditionally stable.

The power amplifier yields a 0.7 dB droop across the 902 to 928 MHz in the ISM band, provides a power to a balanced 50Ω load of up to +13 dBm with a 35 dB lower range, and attains a peak conversion efficiency of 42% at maximum power, including the preamplifier. The 1 dB compression point in the amplifier input-output characteristic almost coincides with the peak power (Figure 10).

Local Oscillator with Quadrature Outputs

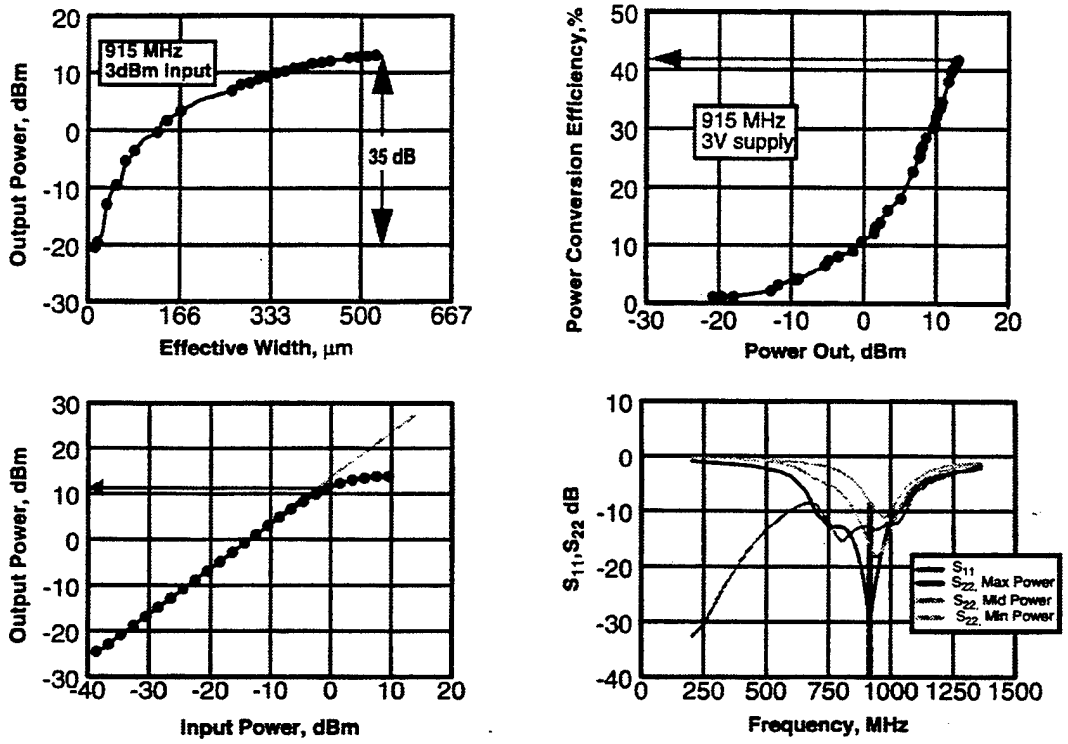


Figure 10: Measured RF Power Amplifier Characteristics

In keeping with the desire for maximum integration, an on-chip fixed-frequency local oscillator was sought which does not require an external resonator. Aside from the challenge of making a 1µm CMOS oscillator operate at all at 900 MHz, there is the additional requirement that the oscillation must be accompanied by a low phase noise, and that the oscillator should produce a large swing to drive the four-FET switch mixer. Incidentally, the last two requirements are in accord with each other, rather than in conflict.

The oscillator core comprises a cross-coupled, common-source FET pair providing a negative resistance to inductor loads (Figure 11). The inductors are fabricated on-chip. Parallel resonance sets in between the inductors and the net FET capacitance, part of which is a drain junction capacitance with a voltage coefficient. This latter is used as a means of coarse frequency tuning, by varying the top-rail voltage with a FET resistor carrying the common-mode current (Figure 12). However, the voltage swing at the oscillator and its current drain will now change with the frequency. The core circuit drains 7 mA from a 3V supply at 915 MHz, and spans the range from 860 to 958 MHz.

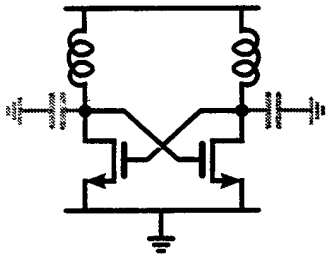


Figure 11: LC Oscillator Core

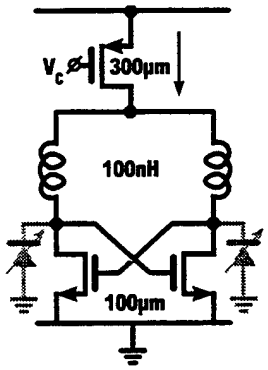


Figure 12: Tuning method

Clipping in the FET characteristics determines the amplitude of oscillation, and this in turn depends on the top-rail voltage. Unlike transistors with a junction at the input which clips by turning ON, the swing at the MOSFET gate may be arbitrarily large, limited only by breakdown. Thus, the oscillator produces a 6.5V ptp differential output at 915 MHz, and for a given transistor noise and Q (=5.5) of the resonant circuit, the corresponding output phase noise is lowered in inverse proportion [11]. An SSB phase noise of –100 dBc/Hz is measured at a 100 kHz offset from the 915 MHz oscillation (Figure 13). The slope of the phase noise vs frequency is 30 dB/decade

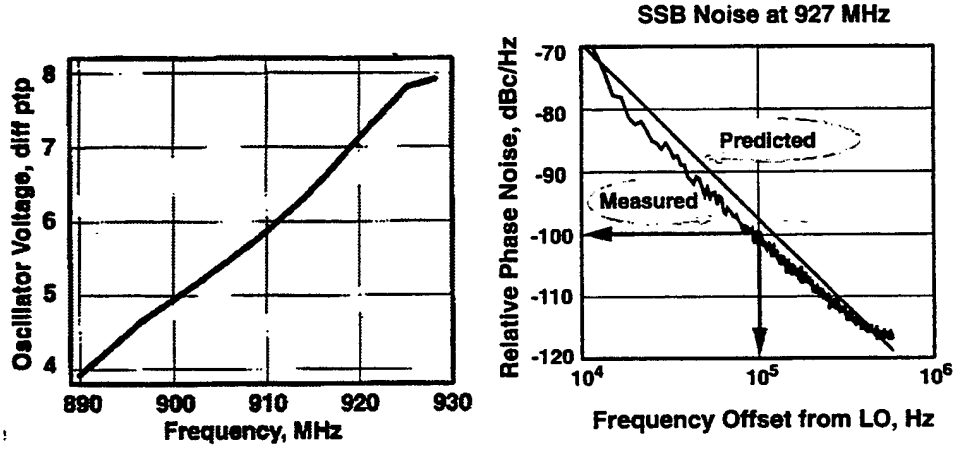


Figure 13: Measured LO characteristics

up to a 1 MHz offset, which shows that MOSFET flicker noise dominates. We are able to predict this very well using Leeson's model [11], and our own measurements of flicker noise in MOSFETs [12, 13].

Single-sideband modulators require quadrature phases of the local oscillator, usually in the form of balanced signals. Quadrature must further be attained with considerable accuracy for adequate sideband suppression (Figure 7). Most local oscillators produce an unbalanced single-phase oscillation, and various methods, which by and large are marginally satisfactory, are used to derive quadrature phases [6]. This LO uses a new topology to inherently produce balanced outputs in precise quadrature (Figure 14) [14]. It consists of two identical oscillators

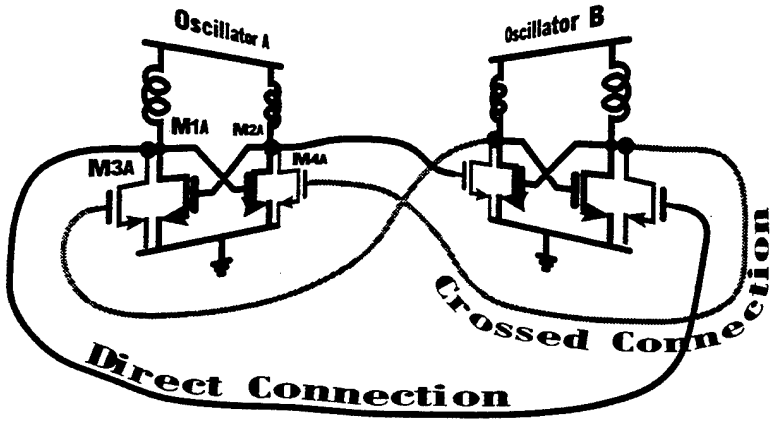


Figure 14: LC Oscillators Synchronized in Quadrature

tightly coupled through FETs so that they injection lock to exactly the same frequency, but the coupling topology suppresses both in-phase and anti-phase synchronization, and permits only synchronization in quadrature with a unique sequence of phases. This negative feedback suppresses errors from quadrature caused by asymmetrical capacitive loading.

The accuracy of quadrature is measured with an on-chip single-sideband upconversion (Figure 15), where it is found that the unwanted sideband is suppressed by more than 46 dB, and LO leakage is 49 dB lower than

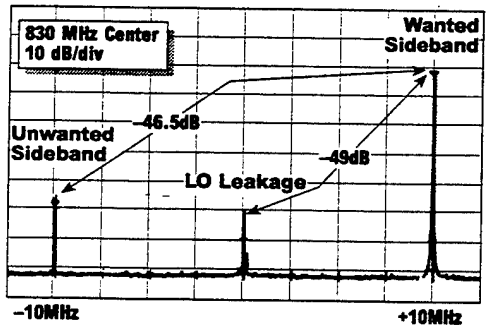


Figure 15: SSB upconversion to measure quadrature accuracy

the wanted sideband. This implies a phase error of less than 1°.

Integrated Transmitter

The various blocks described above have been integrated into a monolithic transmitter, containing all functions from baseband data in to antenna drive at the output (Figure 16) [15]. Overall performance of the transmitter is evaluated, first, by measuring the spectrum of a single tone produced in the ISM band (Figure 17), and, second, by measuring the spectrum with frequency-hopping (Figure 18). In both cases the unwanted sideband and LO leakage are at least 43 dB below the desired parts of the transmitted spectrum.

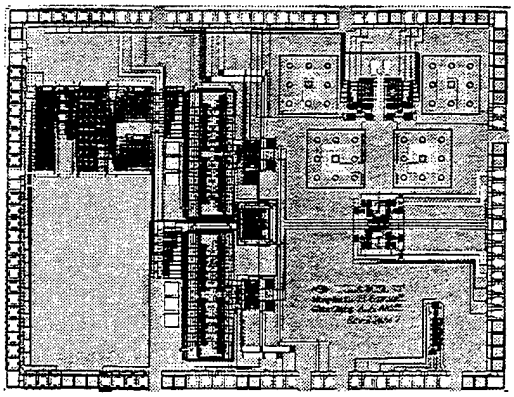


Figure 16: Monolithic Transmitter

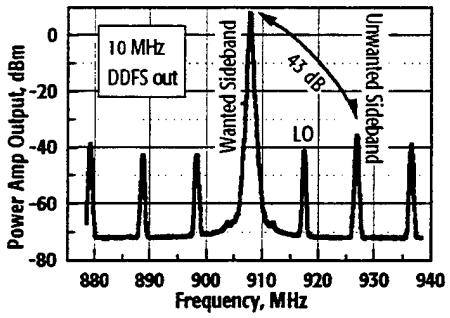


Figure 17: Measured spectrum with single-tone output (med power)

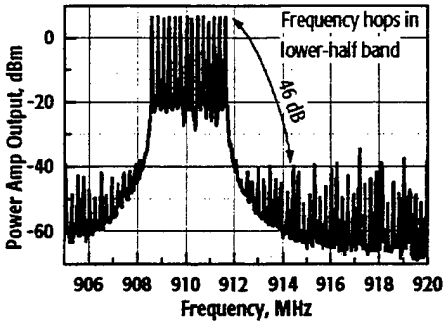


Figure 18: Measured spectrum with frequency hopping (med power)

Technology for Large-Value On-Chip Inductors

An enabling component in realizing RF blocks in the transmitter and receiver is a unique method to realize on-chip spiral inductors, 50 to 100 nH in value, and with a self-resonance beyond 2 GHz. Large-value spiral inductors are conventionally thought to be incompatible with silicon IC technology, because the large parasitic capacitance they inherit through oxide to the semiconducting substrate causes self-

resonance at a few tens or hundreds of MHz. We have overcome this limitation by using a simple, post-process technique that selectively removes the silicon substrate with a gas-phase etchant, leaving the inductor encased in a membrane of oxide attached to the rest of the substrate, but suspended above an air cavity (Figure 19) [16]. In the MOSIS 1-μm CMOS process, this does not require any extra masks, and the etching is carried out on fully fabricated die. The inductor Q of 4 to 5 is limited by the sheet resistance of the 2nd-level Al metallization.

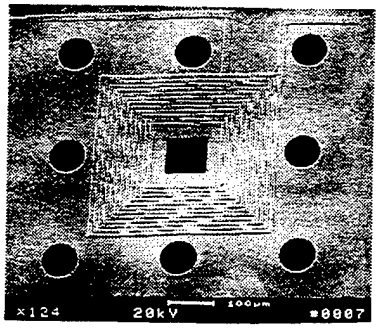


Figure 19: Large spiral inductor suspended over cavity

In many circuits, the larger the available inductor, the lower the bias current required to achieve a certain specification. The main advantage in realizing the inductor on-chip is that no power is wasted in driving the pad, package, and board stray capacitances at radio frequencies which an off-chip, discrete inductor would entail. A simple, three-element model suffices for circuit simulations [16], where the inductance, resistance, and parasitic capacitance to the far-off ground plane may be manually calculated using well-known formulas.

Low-Noise Amplifier and Downconversion Mixer

A receiver front-end which meets the quartet of requirements of low noise, high gain, wide dynamic range, and a good input match to 50Ω at 900 MHz may well be thought to be the greatest design challenge in a 1-μm CMOS implementation. Yet a state-of-the-art circuit has been designed which accomplishes all four of these objectives at a modest current drain. The design relies on a uniquely CMOS approach, and exploits the simplifications implied by direct-conversion [6]. Furthermore, the two components of the front-end, the LNA and mixer, are designed simultaneously.

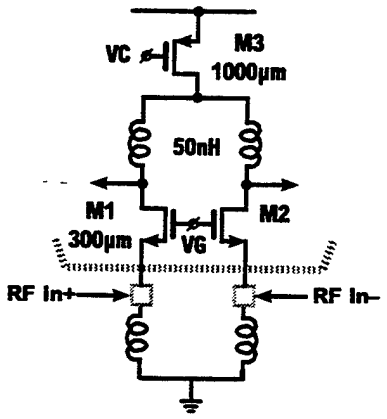


Figure 20: RF Low Noise Amplifier

When a 3 dB noise figure is acceptable in the front-end, it is much simpler to use a common-gate FET input stage whose transconductance is designed to be 1/50 S, than to use the more elaborate inductor matching techniques when the input is applied to the FET gate [1]. A broadband input match now requires the addition of only an off-chip shunt inductor to tune out the pad and FET capacitance at the source. This has been implemented in a quasi-differential circuit (Figure 20) [17]. On-chip inductor loads tune the drains of the FETs to 1 GHz, including the load capacitance of the directly-coupled downconversion mixer (Figure 21), and the Q is sufficiently high that a voltage gain larger than 22 dB is obtained. The input noise figure is set by the thermal noise at 1 GHz in the FET inversion layer, and is about 3 dB. The downconversion mixer (Figure 21) [17] comprises a linear FET voltage-to-current converter connected directly to the LNA output, whose output is commutated by the LO through four FET switches. Two common-source FETs with balanced input and output constitute a very linear transconductor. There is a small conversion loss of 67% when the switches commutate the RF current fully, and balanced FET triode-region loads set the overall mixer gain.

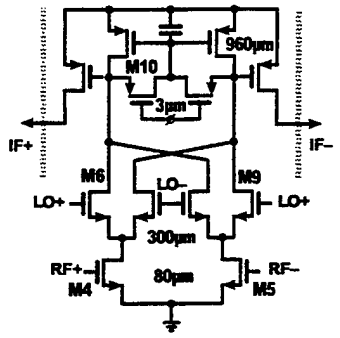


Figure 21: Downconversion Mixer

The measured data (Figure 22) verifies a satisfactory design. In a direct-conversion receiver, the appropriate noise figure is measured with a double-sideband input [18], because both sidebands around the LO carry useful signal energy. The LNA and *one* mixer take 8 mA from a 3V supply when driving an on-chip load. The overall mixer voltage gain is about –3 dB, and it makes a negligible contribution to the overall noise figure. A 0 dBm LO drive is necessary for complete commutation of the mixer switches, and lowest overall noise figure.

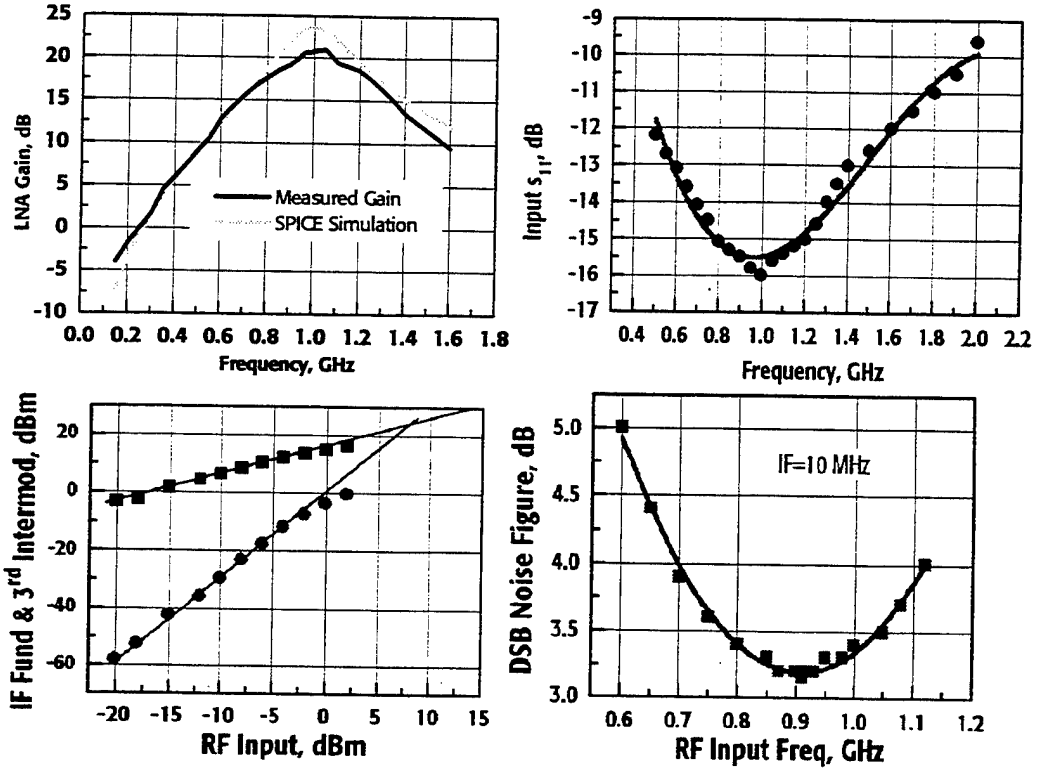


Figure 22: Measured data for LNA-mixer combination

Flicker noise in the baseband section at the mixer output will degrade the noise figure by another 3 dB at 160 kHz, where the peak of the downconverted spectrum lies. Although DC offset at the mixer output may be nulled out in an FSK receiver [6], if the offset varies substantially with LO hopping frequency it becomes difficult to distinguish it from the received symbol energy.

Channel-Select Lowpass Filter

An active on-chip lowpass filter with a sharp cutoff at 230 kHz selects the desired channel, and attenuates all other users within and nearby the ISM band. The filter must itself be low noise so as not to degrade the overall receiver noise figure, it must have wide dynamic range so that out-of-band interferers do not create in-band intermodulation, and it should preferably provide some baseband voltage gain to overcome the input-referred noise of the subsequent limiting amplifier. Receiver linearity remains important until the unwanted signals have been filtered,

and thereafter the received signal is subject to the clipping nonlinearity of the limiting amplifier.

The filter is composed of a decimating switched-capacitor cascade, which progressively amplifies and filters to maintain the largest dynamic range [19]. The downconverted signal is first filtered by a 2nd-order Butterworth section sampling at 57 MHz, which guarantees a 50 dB attenuation at 14 MHz, then decimated by 4 (14.25 MHz clock) into a 6th-order elliptic filter with a 230 kHz cutoff frequency, and the stopband edge at 320 kHz (Figure 23). This partition controls the capacitor spread, which nevertheless is 108:1, while maintaining a well-controlled

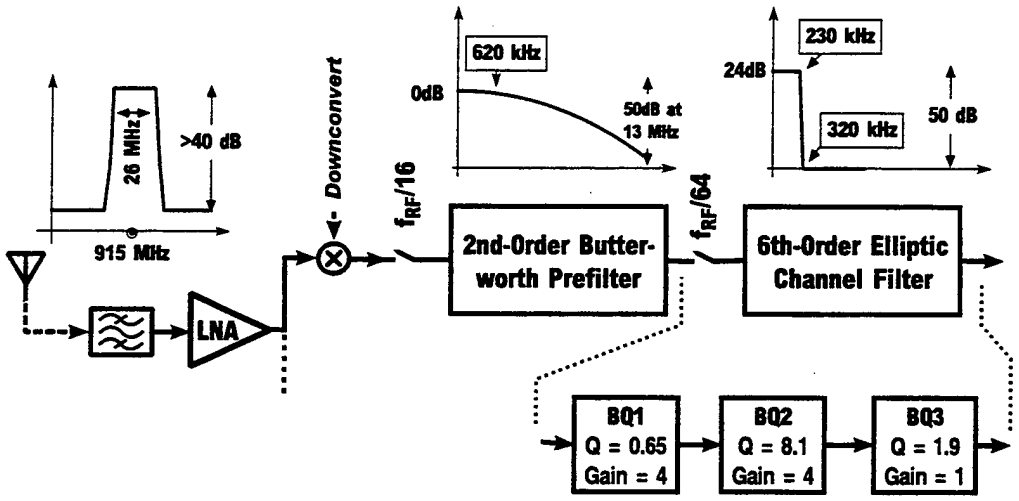


Figure 23: Channel-Select Filter Architecture

stopband to 57 MHz. The unit capacitor size, and therefore the power dissipation, is to a large extent set by the specification on filter input noise.

The measured filter response is very close to what is expected. When consuming 4.6 mA from 3.3V, the input-referred noise in the passband is 70 nV/√Hz. This is dominated by the first stage sampling capacitor, and may be scaled down by redistributing the gain. The IP3 is measured for the case when two large input signals in the stopband create intermodulation distortion in the passband. IP2 characterizes envelope-detection nonlinearity, when an AM signal in the stopband may create detected products lying in the passband. Both measurements bear out that the filter will not become a bottleneck to overall receiver linearity.

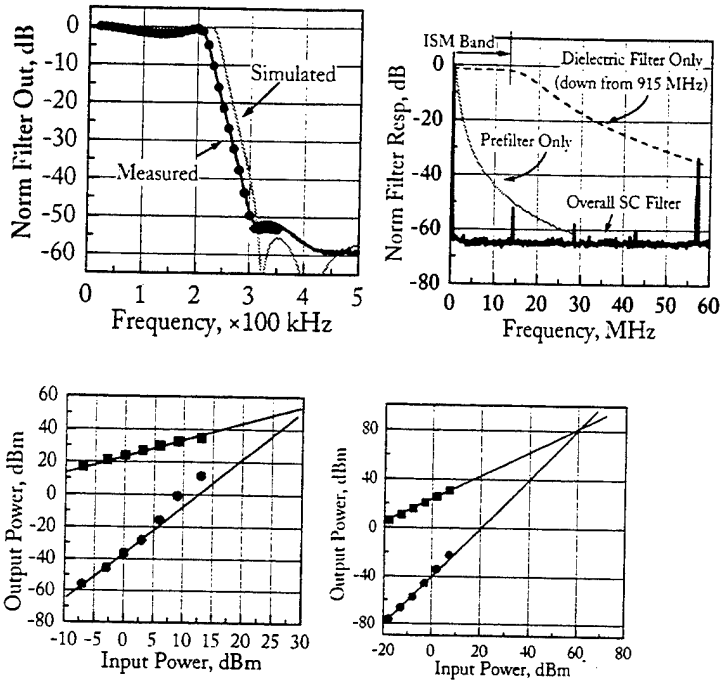


Figure 24: Measurements on Channel-Select Filter

Limiting Amplifier and Detector

A limiting amplifier with 84 dB gain is implemented by a cascade of 7 identical clipping differential pairs. This operates on the selected baseband signal, and so need only have roughly a 1 MHz bandwidth. Rectified taps along this chain provide a successive-detection logarithmic measurement of the received signal strength (Figure 25) [20].

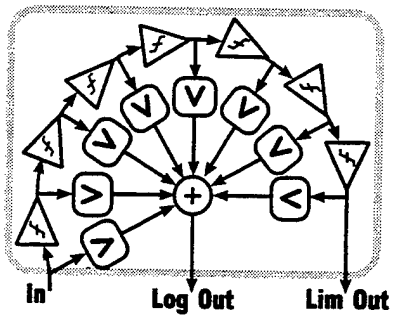


Figure 25: Limiting Amplifier with Log Signal-Strength Output

A capacitively-decoupled DC feedback loop is applied between the output and input of the 2nd

stage, to suppress DC offset from dominating the limiting action. With this feedback, the circuit tolerates a ±100 mV offset on the input signal without loss of dynamic range. Typical of a sensitive limiting amplifier, in the absence of an input this circuit limits on its own noise, which is 50μV at the input in a 300 kHz bandwidth. The logarithmic output is 1 dB accurate over an 80 dB range of the input across the commercial temperature range. The circuit drains only 1 mA from 3.3 V.

A digital detector accompanies the limiting amplifier (Figure 26) [21]. This oversamples the limited output of the downconverted desired channel by 64×, then correlates with locally generated 1b *sin* and *cos* waveforms (square waves) at the expected frequency offset. The correlations are integrated over 64 clock cycles, corresponding to one symbol, and a decision is made based on an estimate of the energy of the largest correlation. Various simplifications are made to retain 1b signal processing in the portions of the circuit clocking at high speed. When preceded by two limiting amplifier channels, and accompanied by a digital oscillator for the reference sinewaves, this circuit only dissipates 5 mW from 3.3 V. Yet with the limiting amplifier preceding it, the circuit operates at an acceptable error-rate for a 56μV rms input, and can tolerate inputs at least 82 dB larger than this.

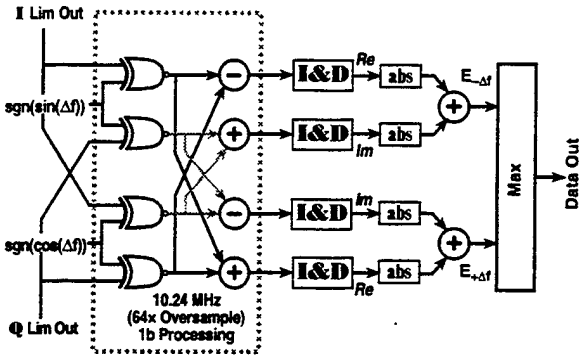


Figure 26: Correlating digital binary-FSK detector

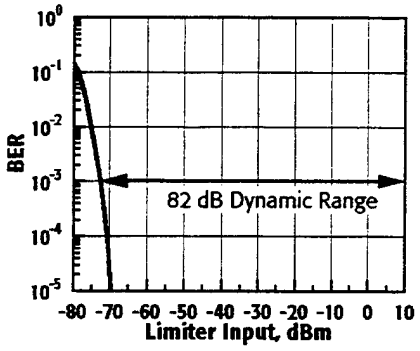


Figure 27: Measured FSK sensitivity of limiting amplifier and detector

The digital circuits for timing acquisition and frequency synchronization are described elsewhere [22].

A fully integrated receiver, including all functions from antenna to baseband data output, is now being evaluated (Figure 28).

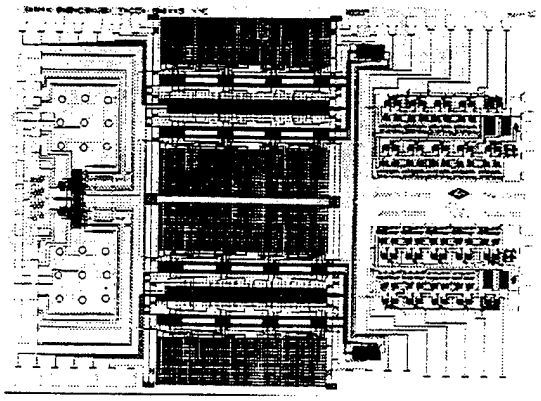


Figure 28: Integrated Receiver IC

REFERENCES

[1] A. A. Abidi, "Low-Power Radio-Frequency ICs for Portable Communications," *Proc. of the IEEE*, vol. 83, no. 4, pp. 544-569, 1995.

[2] I. A. W. Vance, "Fully Integrated Radio Paging Receiver," *IEE Proceedings*, vol. 129, Part F, no. 1, pp. 2-6, 1982.

[3] D. W. H. Calder, "Audio Frequency Gyrator Filters for an Integrated Radio Paging Receiver," presented at *Int'l Conf on Mobile Radio Systems & Techniques*, York, UK, pp. 21-26, 1984.

[4] P. E. Chadwick, "Analogue VLSI-applications of the technique," presented at *IEE Colloquium on 'Advances in Analogue VLSI'*, London, pp. 1/1-3, 1991.

[5] J. Min, A. Rofougaran, H. Samueli, and A. A. Abidi, "An All-CMOS Architecture for a Low-Power Frequency-Hopped 900 MHz Spread-Spectrum Transceiver," presented at *Custom IC Conf.*, San Diego, CA, pp. 379-382, 1994.

[6] A. A. Abidi, "Direct-Conversion Radio Transceivers for Digital Communication," *IEEE J. of Solid-State Circuits*, vol. 30, no. 12, pp. 1399-1410, 1995.

[7] H. T. Nicholas and H. Samueli, "A 150-MHz Direct Digital Frequency Synthesizer in 1.25μm CMOS with -90dBc Spurious Response," *IEEE J. of Solid State Circuits*, vol. 26, pp. 1959-1969, 1991.

[8] G. Chang, A. Rofougaran, M. K. Ku, A. A. Abidi, and H. Samueli, "A Low-Power CMOS Digitally Synthesized 0-13 MHz Agile Sinewave Generator," presented at *Int'l Solid State Circuits Conf.*, San Francisco, pp. 32-33, 1994.

[9] R. C. V. Macario and I. D. Mejallie, "The Phasing Method for Sideband Selection in Broadcast Receivers," *EBU Review (Technical Part)*, no. 181, pp. 119-125, 1980.

[10] M. Rofougaran, A. Rofougaran, C. Olgaard, and A. A. Abidi, "A 900 MHz CMOS RF Power Amplifier with Programmable Output," presented at *Symp. on VLSI Circuits*, Honolulu, pp. 133-134, 1994.

[11] D. B. Leeson, "A Simple Model of Feedback Oscillator Noise Spectrum," *Proc. of IEEE*, vol. 54, no. 2, pp. 329-330, 1966.

[12] A. A. Abidi, J. Chang, C. R. Viswanathan, J. Wikstrom, and J. Wu, "Uniformity in Flicker Noise Characteristics of CMOS IC Technologies," presented at *Ninth International Conference on Noise in Physical Systems*, Montreal, pp. 469-472, 1987.

[13] J. Chang, A. A. Abidi, and C. R. Viswanathan, "Flicker Noise in CMOS Transistors from Subthreshold to Strong Inversion at Various Temperatures," *IEEE Trans. on Electron Devices*, vol. 41, no. 11, pp. 1965-1971, 1994.

[14] A. Rofougaran, J. Rael, M. Rofougaran, and A. A. Abidi, "A 900 MHz CMOS LC Oscillator with Quadrature Outputs," presented at *Int'l Solid State Circuits Conf.*, San Francisco, CA, 1996.

[15] A. Rofougaran, G. Chang, J. J. Rael, M. Rofougaran, S. Khorram, M.-K. Ku, E. Roth, A. A. Abidi, and H. Samueli, "A 900 MHz CMOS Frequency-

A Low-Power Baseband Receiver IC for Frequency-Hopped Spread Spectrum Applications

H.-C. Liu, J. Min, and H. Samueli

IEEE Journal of Solid-State Circuits
vol. 31, no. 3, pp. 384-394
March 1996

A Low-Power Baseband Receiver IC for Frequency-Hopped Spread Spectrum Communications

Huan-Chang Liu, *Student Member, IEEE*, Jonathan S. Min, and Henry Samueli, *Member, IEEE*

Abstract—**A baseband receiver IC which will be incorporated into a low-power frequency-hopped spread spectrum (FH/SS) transceiver for 902–928 MHz ISM band applications is presented. The chip performs noncoherent binary/quaternary frequency shift keying (FSK) demodulation, equal-gain diversity combining of dual antenna branches, and symbol and frequency synchronization. The chip also accommodates variable data rates from 2 to 160 kb/s, programmable hop rates, and tunable bandwidth loop filters. The core area of the 1-μm CMOS chip is 3.9 mm $\times$ 3.9 mm with a power consumption of 4.5 mW at 10 MHz from a 3-V supply. A baseband transceiver system utilizing this receiver chip for the prototype handset to demonstrate a point-to-point communication link is also described. Two XILINX FPGA chips were used to implement the remainder of the baseband transceiver functions, including frequency control logic for FSK modulation, acquisition control, data framing, symbol interleaving and deinterleaving, and interface control for data and voice.**

I. Introduction

SPREAD spectrum techniques have recently attracted much attention for wireless communications applications. The frequency-hopped spread spectrum (FH/SS) technique is especially suitable for a low-power handheld communications device [1]. The frequency shift keying (FSK) modulator and demodulator section of the baseband transceiver presented in this paper are part of an all-CMOS monolithic FH/SS transceiver which is currently under development at the University of California at Los Angeles [2]. The FH/SS transceiver is intended for use in a wide variety of portable wireless applications in the unlicensed 902–928 MHz industrial, scientific, and medical (ISM) band. Low power and robustness are two key requirements for a personal communications device. A direct-conversion receiver architecture from RF to baseband (Fig. 1) reduces the overall power consumption by eliminating off-chip filters. Fast frequency hopping is implemented using a direct digital frequency synthesis technique [3]. A single sideband (SSB) modulation technique, in which both a quadrature direct digital frequency synthesizer (DDFS) and a quadrature local oscillator are required, is utilized to suppress the unwanted image signals without resorting to often difficult-to-implement image-reject filters when a baseband FSK signal is upconverted to RF.

Manuscript received October 22, 1995; revised December 13, 1995. This work was supported by ARPA/ETO under Contract DAA B07-93C-C501.

The authors are with Integrated Circuits and Systems Laboratory, Electrical Engineering Department, University of California, Los Angeles, CA 90095-1594 USA.

Publisher Item Identifier S 0018-9200(96)02458-4.

The FH/SS technique provides frequency diversity over multipath fading channels. Fast hopping, however, mandates a noncoherent FSK detector. To guarantee robust digital data transmission, the receiver also incorporates dual antenna diversity. Consequently, the demodulator needs to combine the FSK tone energies of the two receive branches. In order to efficiently use the entire 26 MHz transmission bandwidth, a quadrature demodulation technique is used. This requires two channels (I & Q) for each branch. The automatic gain control (AGC) function is implemented using a combination of a lowpass filter (LPF) and a hard limiter, each capable of a wide dynamic range (>80 dB). The LPF is implemented using a switched-capacitor technique for its precise tuning capability. The hard limiter works as a 1-b quantizer, and makes the received signal independent of level variations. Due to odd harmonics produced by hard-limiting the sine waves, this approach limits the modulation scheme to binary FSK for real signal detection and to quaternary FSK for complex signal detection. However, no analog-to-digital converters or linear variable gain amplifiers are required, which significantly reduces the overall hardware complexity.

II. Baseband Transceiver System

A. Baseband Transmitter

The baseband transmitter is quite simple in this system, since both FSK and FH are achieved using the DDFS. Only the frequency control logic and the DDFS/digital-to-analog converter (DAC) combination are required. The frequency control logic consists of an encoder that translates the input symbol into an appropriate quaternary FSK frequency and a memory that stores the frequencies which correspond to the hopping patterns. These two frequencies are summed at every baud cycle and fed into the frequency control word of the DDFS for proper tone generation. The use of the DDFS results in continuous phase changes between different frequencies which makes the transmitted FSK signal a form of continuous-phase modulation (CPM). Although the phase continuity of the CPM process cannot be utilized at the receiver since the channel introduces a phase jump between hops, low spectral sidelobes of the transmitter CPM spectra are still beneficial for multiuser radio scenarios.

B. Baseband Receiver

There are many ways to detect FSK signals. However, both the direct-conversion and the frequency hopping fea-

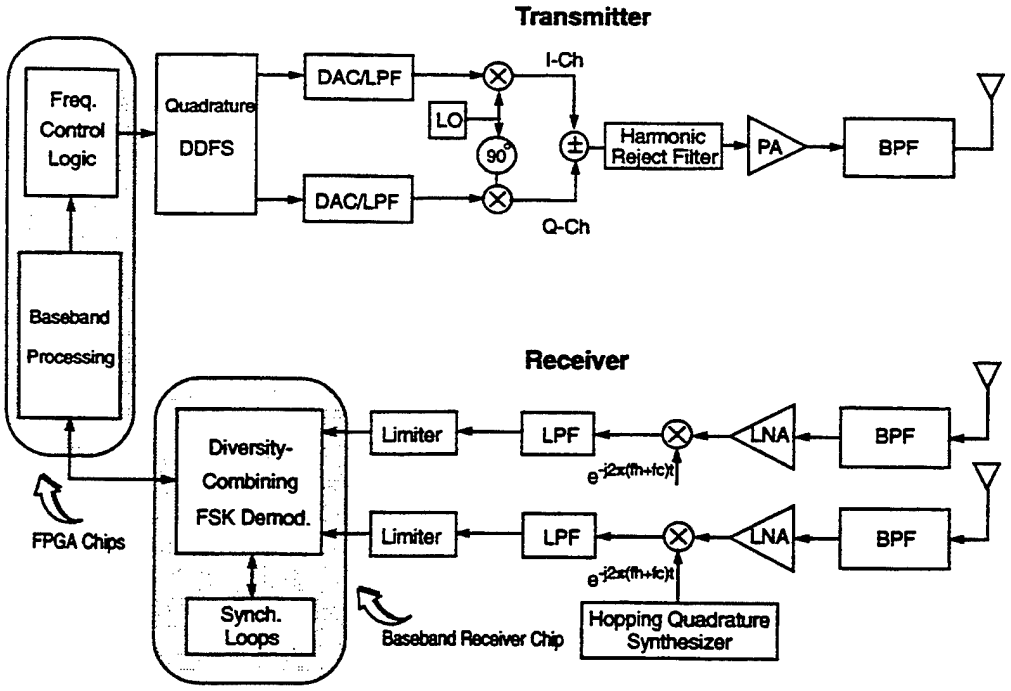


Fig. 1. Overall FH/SS transceiver system block diagram.

tures of this receiver architecture constrain the choice of the FSK detector. For a direct-conversion receiver, a quadrature baseband detector is required. Phase discontinuity due to frequency hopping excludes some types of FSK detectors. For a conventional superheterodyne receiver, IF FM detectors such as a limiter-discriminator are typically used; however, this approach cannot be applied to our direct-conversion receiver. Simple flip-flop detectors or zero-crossing detectors [4] commonly used in pagers with large modulation indices are not suitable for new digital wireless systems where bandwidth-efficient modulation implies only one signal cycle per baud. Recently, new baseband detectors for binary FSK for the DECT and CT2 wireless standards have been proposed. In the approach followed in [5], the received I-Q signals are translated into corresponding magnitude and phase, and the difference between two consecutive phases is detected at the digital detector. Not only does this differential scheme require complex signal processing, but it is limited only to low hop-rate systems such as a slow frequency-hopped time-division multiple access (TDMA) system where the radio hops to a new frequency every frame. Because the first bit of each hop is reserved to provide a phase reference to the following bits, too much bandwidth is wasted in a high hop-rate system where phase continuity between hops is not guaranteed. A balanced quotient detector is proposed in [6]; although it may be simple to implement, it is still analog in nature and has certain limitations. For example, clock recovery would be difficult, if not impossible, for a fast frequency-hopped receiver.

It is well-known that the optimum FSK detector is a correlation detector [7]. However, this detector is not often used in practice owing to the complexity of the required circuits. A simple 1-b correlation FSK detector has been designed for use in this frequency-hopped direct-conversion receiver. Not only does this choice eliminate expensive digital multipliers, but it also provides a flexible design which can easily accommodate programmable data rates, and the same hardware can be used to detect both data and synchronization hops. However, the oversampling ratio ($T_{\text{baud}}/T_{\text{clk}}$) must be chosen properly—multiples of four for binary FSK and multiples of eight for quaternary FSK—in order to avoid harmonic aliasing, thus minimizing performance degradation caused by one-bit correlation [8].

With open-loop direct conversion from a 915 MHz RF to baseband, the carrier frequency error could amount to ±100 kHz between the transmitter and receiver assuming ±50 ppm crystal accuracies. The receiver therefore requires a frequency tracking loop as well as a baud time tracking loop. Phase-locked loop (PLL) techniques are most commonly used to correct either the timing error or the frequency offset between the transmitter and receiver. Either analog or digital implementation is possible for a PLL. The speed of the clock is one of the factors that gives an advantage to analog PLL's compared to digital PLL's. If the required clock rate is too high (for example, more than 10 MHz), the loop typically requires an analog voltage-controlled crystal oscillator (VCXO) because a numerically controlled oscillator (NCO) implementation [9] would be impractical. If a DDFS technique is used for carrier generation as in digital IF receivers, a fully digital loop architecture is appropriate. Since a DDFS is used for hopping carrier generation and the baud clock frequency is relatively low (80 kHz) in our system, both time tracking and frequency tracking loops have adopted a fully digital-PLL architecture. For a tracking loop, a first-order loop filter is typically used, thereby making the overall loop second-order. A second-order loop can completely correct for both phase and frequency errors [10]. The first-order digital loop filter block

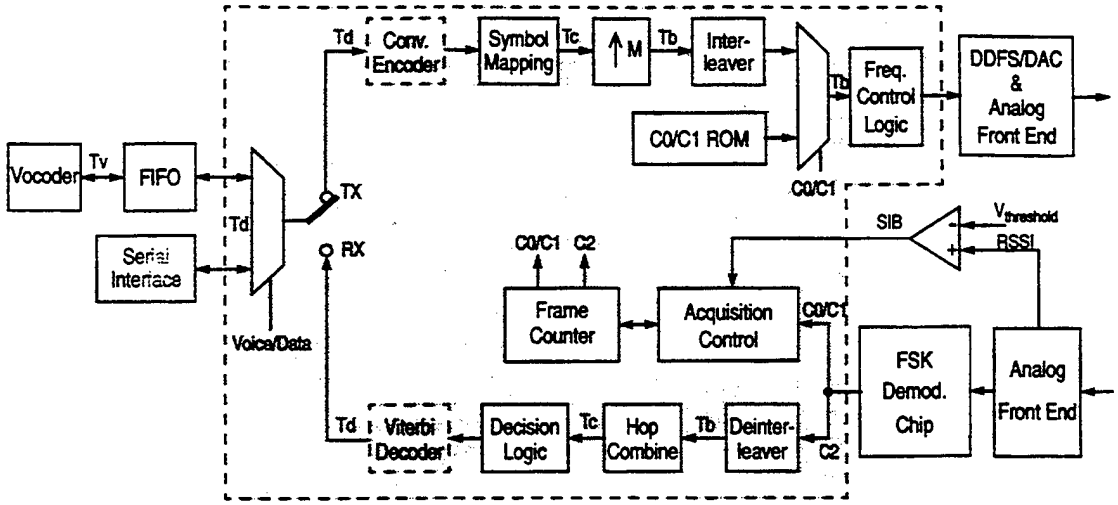


Fig. 2. Baseband processing block diagram.

has two programmable coefficients: the linear term (K0) and the integrator term (K1). K0 typically controls the loop gain while the ratio K0/K1 determines the loop bandwidth.

C. Baseband Processing Blocks

The remaining baseband processing blocks preceding the frequency control logic or following the demodulator include a baseband controller which handles acquisition and frame synchronization in the receive mode, an interleaver/deinterleaver, a hop repeater/combiner, an optional channel coder/decoder, and an interface for voice and data. The overall baseband processing blocks developed for the prototype handset are shown in Fig. 2. The algorithms and architecture for the proposed acquisition protocol are based on a simple time division duplexing (TDD) frame structure, which consists of a pilot tone (C0), a frame ID (C1), and actual data (C2). The acquisition process is accomplished by means of frame synchronization, based on coarse energy detection of the received signal and pattern matching of the frame ID (*Word Sync*). For synchronization, the slave handset first listens for the pilot tone of the frame, which is broadcast by the master at a pre-assigned acquisition carrier frequency. The receiver signal strength indicator (RSSI) output, which is generated by a cascade of logarithm amplifiers [11] after the channel-select filter, is then compared to a programmable energy threshold value to produce either high or low output. This signal indicator bit (SIB) indicates whether a signal is present at a particular acquisition frequency or not. Once the bit goes high, frequency acquisition takes place by sweeping the frequency of the DDFS in a stepsize of $F_{\text{tone}}/4$ (20 kHz). The frequency sweeping is required since the worst possible frequency error (±100 kHz) may be more than what the frequency tracking loop can handle. The five-bit correlation value of a predetermined tone ($+F_{\text{tone}}$) is evaluated to indicate a coarse frequency lock. After coarse frequency acquisition, the rest of the pilot tone slots are used to detect a signal transition from $+F_{\text{tone}}$ to $-F_{\text{tone}}$ within a T_{baud} uncertainty. A 21-digit binary Barker code [12] is employed as the C1 ID code for frame synchronization. It has an autocorrelation characteristic in which its autocorrelation value peaks when the reference and received codes align; otherwise, its value is quite small (+1), often negative numbers. Rather than having 21 independent digits, its pattern consists of three subsets of a seven-digit code $\{1\ \ 1\ \ 1\ \ -1\ \ -1\ \ 1\ \ -1\}$: the first two sets have a positive polarity and the third one has a negative polarity. Therefore, the received data should be matched with the unique ID pattern before frame synchronization is declared. Once *frame sync* is declared, the receiver starts demodulating data. The hopping code synchronization relies on the fact that once the frame is synchronized, a predefined frequency hopping pattern is repeated every frame. Thus, no extra PN code acquisition is required. Compared to a direct-sequence (DS) system where the receiver typically has to go through a complicated serial PN search process, this method makes FH acquisition much faster.

The proposed transceiver can handle various data rates between 2–160 kb/s. When the data rate is lower than the modulation rate of 80 kbaud for quaternary FSK, the data may be repeated in symbol format. This repeat code provides extra diversity over a Rayleigh fading channel when independent hops of the same information are combined after deinterleaving. Since our system is slow FH with eight symbols per hop, a convolutional symbol interleaver/deinterleaver whose size is 8 × 40 is inserted between the modulator/demodulator and hop repeater/combiner. A convolutional interleaver is chosen over a block interleaver, since its delay is half the delay of the block interleaver and the hop phase jump is randomized, unlike in the block interleaver. The column size of 40 symbols for the interleaver provides enough separation for the proposed rate-1/2 Viterbi decoder to operate in an optimal condition [13].

III. Baseband Reciever Chip Implementation

The shaded area of Fig. 3 shows the building blocks implemented in this baseband receiver chip. Since the target application is for a handheld transceiver, both low power and

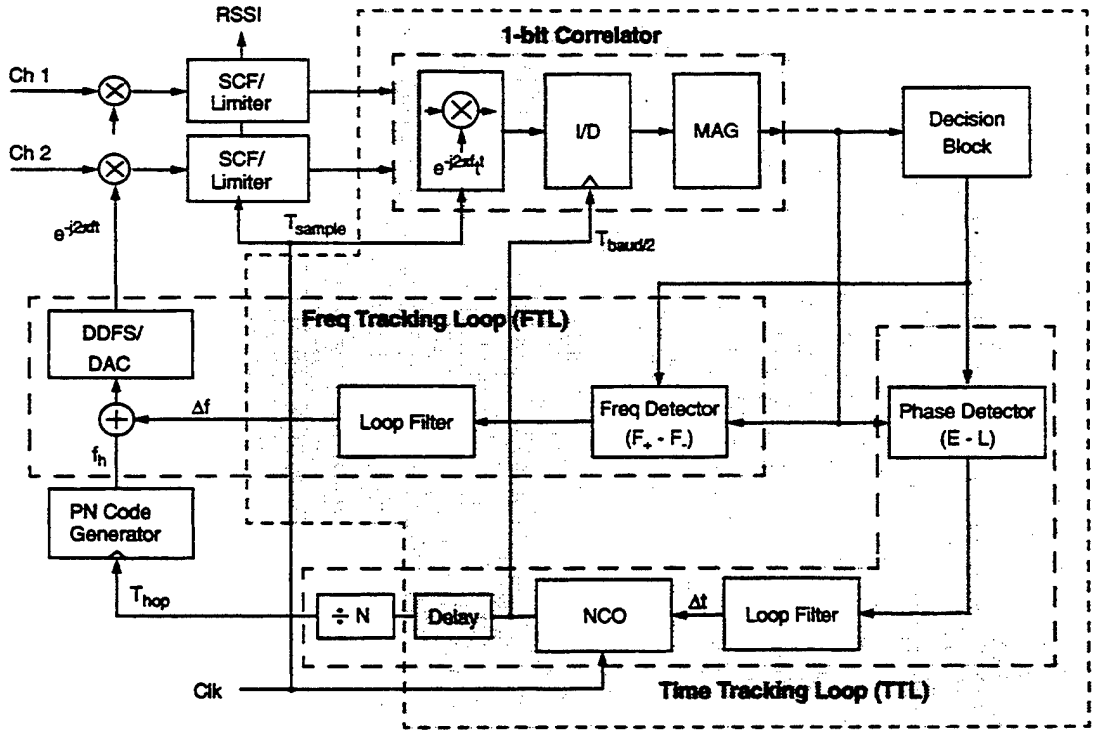


Fig. 3. Baseband receiver chip block diagram.

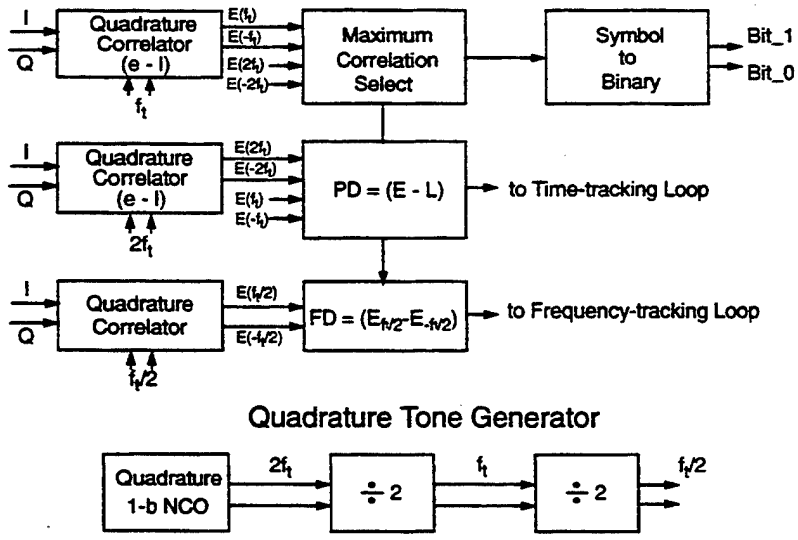


Fig. 4. Overall digital demodulator architecture.

small area are important considerations. Several techniques are applied in different levels of the design to achieve this goal. Since the input is only one bit, the local tone generator is also implemented using a 1-b-output NCO, rather than a high-precision DDFS [14]. The demodulator supports both binary and quaternary FSK signaling. The binary FSK format consists of the tones $\{+F_{\text{tone}}, -F_{\text{tone}}\}$, while the quaternary FSK format contains the tones $\{-2F_{\text{tone}}, -F_{\text{tone}}, +F_{\text{tone}}, +2F_{\text{tone}}\}$. Thus, complex $\pm F_{\text{tone}}$ and $\pm 2F_{\text{tone}}$ detectors are required. In addition, a $\pm F_{\text{tone}}/2$ correlator is used in the frequency tracking loop. The quadrature NCO generates a 1-b I-Q $2F_{\text{tone}}$ square-wave and the other reference tones, F_{tone} and $F_{\text{tone}}/2$, are easily produced by simple binary division of the NCO tone (Fig. 4).

A noncoherent demodulator requires a magnitude calculation unit, which is conventionally implemented with a pair of squaring multipliers followed by a summing node. In our architecture, an absolute value addition block (Fig. 5) replaces this squaring block. Thus, a truly multiplierless noncoherent FSK demodulator has been implemented with little performance degradation. Symbol timing must be derived from the received signal in order to synchronously sample the output

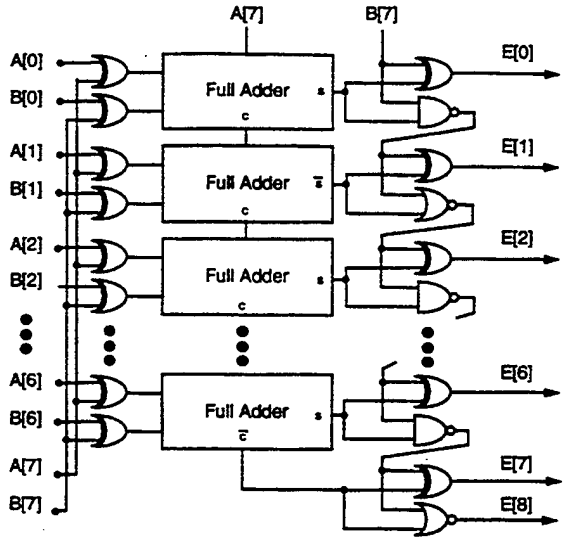


Fig. 5. Absolute-value summing node block diagram.

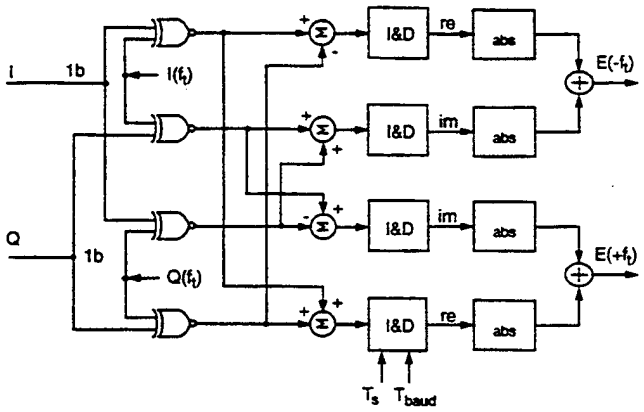


Fig. 6. Multiplierless quadrature 1-b FSK detector block diagram.

of the correlators at the proper time instant. An early-late gate synchronizer has been designed to control the clock timing of the integrate-and-dump (I/D) units in the tone detector. The PN code generator provides the frequency hopping pattern which is fine-tuned by the output of the loop filter, and controls the frequency word of the DDFS/DAC block. Since in our system the hop rate is slower than the baud rate, the hop time synchronization is performed in an open-loop fashion, aligning the hop time with the baud time plus a programmable delay to compensate for the delay between the two loops. Circuit parallelism is employed in the blocks operating at a high clock rate to eliminate the power consumption of time multiplexing and control. The supply voltage can also be scaled down to further reduce the power consumption. On the other hand, time-multiplexed implementation is applied on the blocks with a low clock rate to minimize the chip area. A robust, gated clocking scheme is used for registers with a low access rate to further reduce the power consumption. In order to support low data rate operations, static D flip-flops are used in the entire chip. The data race problem is avoided by distributing the clock signals in the opposite direction to the flow of data. The transistor sizes are optimized with HSPICE simulations over process variations. Minimum-sized devices are used wherever possible. Transmission-gate logic is also used where driving capability is not a problem to minimize the chip area. For a low-power CMOS circuit design, the logic functions have been reordered and simplified to reduce the number of gates, as shown, for example, in the absolute value addition block in Fig. 5.

A. FSK Detector

The FSK detector consists of complex data mixers, an integrate-and-dump block, a magnitude calculation unit, and hard-decision logic (Fig. 6). Many architectural optimization techniques have been incorporated into the design. First, the front-end mixer uses an exclusive-NOR gate as a 1-b multiplier. Note that FSK demodulation requires only frequency information, not amplitude or phase. The input signal is oversampled at the clock rate of the switched-capacitor LPF, which has been chosen to meet the oversampling ratio constraint ($N = 88$ for our prototype). However, this rate is immediately decimated down to twice the baud rate following the integrate-and-dump block. The integration duration is half the baud time in order to generate the early and late signals for the clock recovery loop. The correlator output can also be obtained by adding these two portions together without the requirement of extra integration hardware. A 7-b transmission gate carry ripple adder with saturation logic is used in the accumulator portion of the integrate-and-dump circuit to avoid overflow which may occur when symbol synchronization has not yet been achieved. To realize dual antenna diversity, the tone energies of the two branches are combined equally after correlation. By forcing a multiplexer pair to pass the correlator outputs of $\pm F_{\text{tone}}$, the hard decision circuitry can easily switch between quaternary and binary FSK formats. A set of 6-b soft decision outputs is also provided for the baseband processing block.

Both synchronization loops are based on a data transition scheme. Therefore, a transition detection block with two kinds of indications is required. One is any symbol transition between the signal sets and the other is a transition from $+F_{\text{tone}}$ to $-F_{\text{tone}}$ or vice versa. These two transition indication outputs are used in the time tracking loop and the frequency tracking loop, respectively. A trade-off between parallel and time-multiplexed implementations was made in order to optimize the power and area. The correlator block with a higher clock rate dominates the power consumption. Therefore, parallelism is adopted to minimize extra power in the time multiplexing and control as shown in Fig. 7. The voltage scaling capability of parallelism can further reduce the power consumption [15]. The blocks after the I/D operate at a much lower rate, so the power consumption is not a critical concern. A time-multiplexed architecture is then chosen to minimize the silicon area.

B. Time Tracking Loop

The time tracking loop consists of a phase detector (PD), a loop filter, and an NCO (Fig. 8). An early–late phase detection scheme is used to generate a phase discriminant.

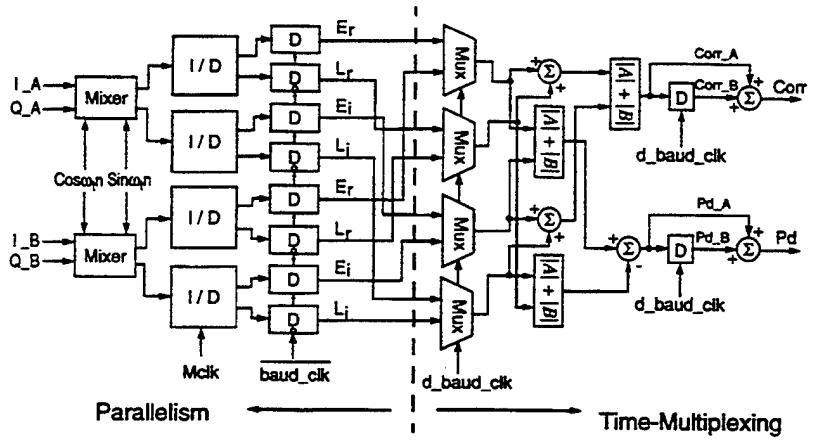


Fig. 7. Detailed quadrature correlator block diagram.

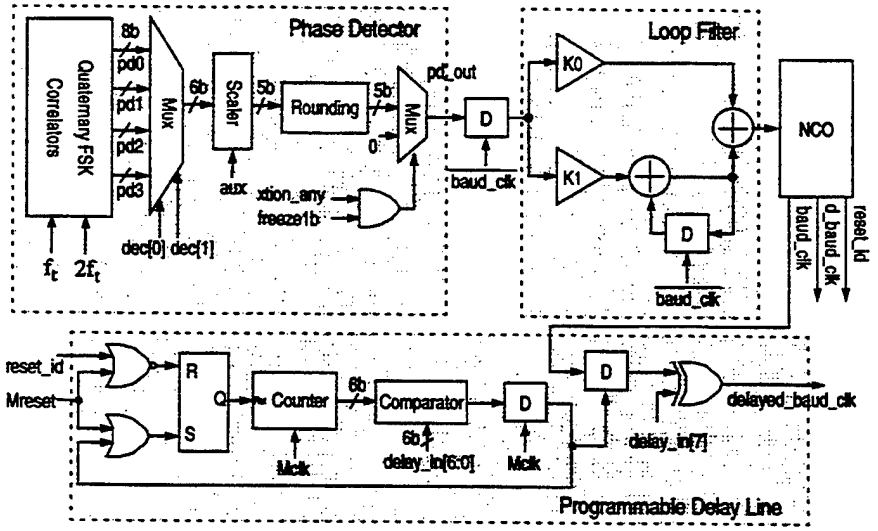


Fig. 8. Time tracking loop block diagram.

The correlation energy of the first-half baud minus that of the second-half baud is detected for each tone and the output from one of the four tones (pd0–pd3), chosen by the hard decision bits, is fed into the loop filter. When only one antenna branch is enabled, a scaler is used before the rounding device to guarantee the maximum dynamic range. The rounding device rounds a 5-b input to a 4-b PD output with one guard-bit at the MSB to prevent an overflow condition. When the data is not changing, the output of the PD should theoretically be zero. Therefore, the phase detector output is enabled only when there is a symbol transition.

A first-order loop filter with programmable powers-of-two coefficients is used to handle various tracking ranges depending on the required clock jitter of the overall system. The loop architecture is simplified by replacing an expensive digital multiplier with an add-and-shift operator and also by limiting the number of coefficient choices to only eight. Saturation circuits are used after the adders in the loop filter to ensure the desired result without overflow. Both the oversampling rate and loop filter coefficients are designed to be fully programmable to accommodate different applications. The output of the loop filter feeds a 24-b NCO, which can provide a 0.5 Hz tuning resolution at a 10 MHz clock rate. The NCO generates the baud clock (*baud_clk*) and a reset signal (*reset_id*) to control the operation of the I/D, and a delayed version of baud clock (*d_baud_clk*) to control the time multiplexing of the correlator. The programmable delay for the hop time synchronization is designed as a transition delay device, rather than a chain of delay registers. It utilizes the periodic and low frequency properties of the baud clock. An exclusive-OR gate provides a half baud clock cycle delay according to the MSB of the *delay_in* control signal. Only one counter and one comparator are needed in this implementation.

C. Frequency Tracking Loop

For the frequency tracking loop (FTL), a frequency detector (FD) replaces a phase detector and a high-precision DDFS is used instead of a 1-b-output NCO (Fig. 9). The difference between the two correlation energies at $+F_{\text{tone}}/2$ and $-F_{\text{tone}}/2$ is used as the baseline frequency discriminator.

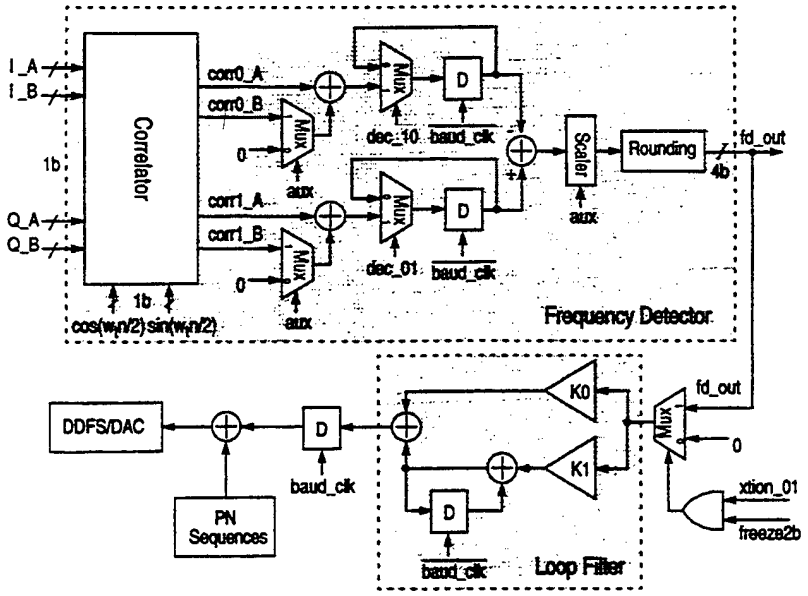


Fig. 9. Frequency tracking loop block diagram.

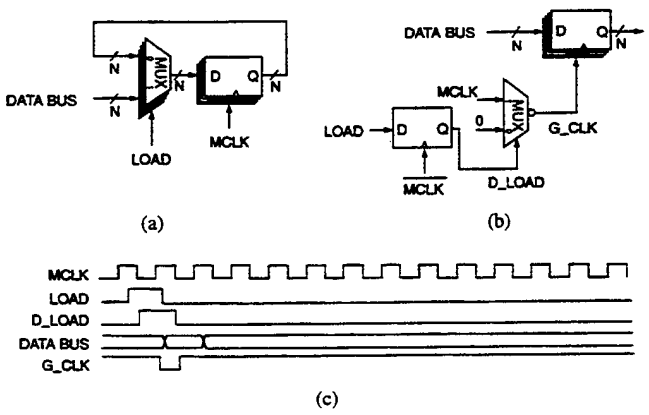


Fig. 10. Gated clocking scheme.

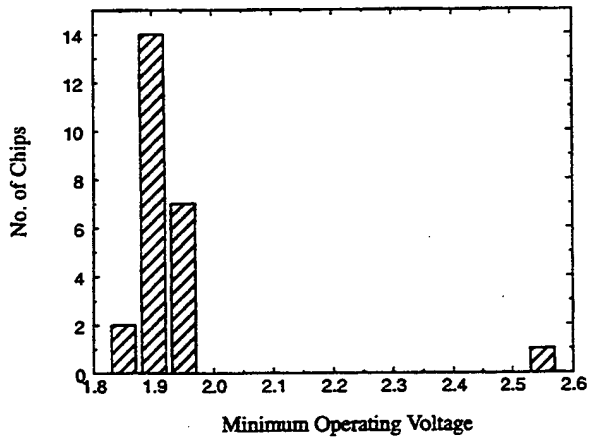


Fig. 11. Minimum operating voltage at 7.04 MHz, $N = 88$.

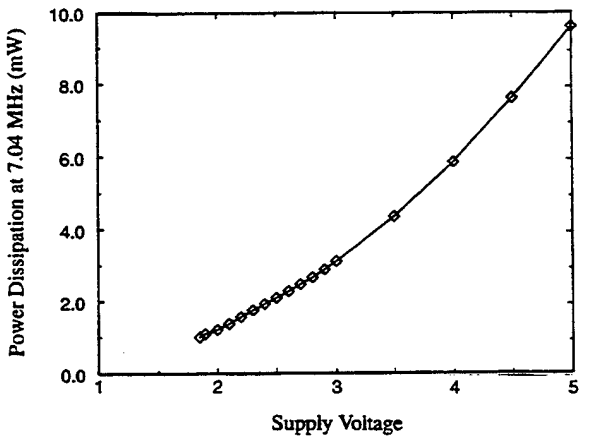


Fig. 12. Power dissipation versus V_{dd} scaling at 7.04 MHz.

When a signal with either $+F_{\text{tone}}$ or $-F_{\text{tone}}$ is received, its correlation value is stored and compared with the other correlation value. The 4-b FD output is then fed into the loop filter only when a transition from $+F_{\text{tone}}$ to $-F_{\text{tone}}$ or vice versa occurs. This scheme simplifies the frequency error detection block significantly and thus justifies its usage in our FTL.

The loop filter has a 24-b internal wordlength and the input frequency control word of the DDFS has been chosen to be 24 bits wide. Though simulations show a 20-b implementation would have been sufficient, this choice fully utilizes the precision of the loop filter output and guarantees a 5 Hz frequency tuning capability at a DDFS clock rate of 76 MHz. Statistically speaking, the loop is activated only in one out of eight clocks due to the data transition restriction in our FD scheme, thus slowing down the frequency tracking process. To speed up the tracking process, the difference between two correlation values at $+3F_{\text{tone}}/2$ and $-3F_{\text{tone}}/2$ can also be used; however, this scheme requires more complex circuits, especially since the $3F_{\text{tone}}/2$ tone cannot be generated by simple binary division from the $2F_{\text{tone}}$ reference tone. Thus,

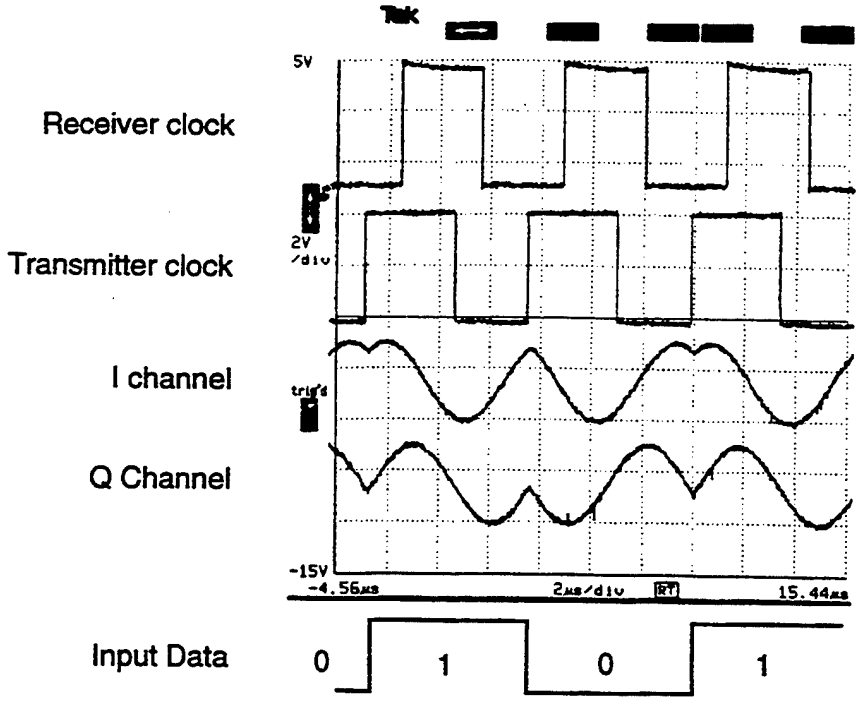


Fig. 13. Plot of the baseband binary FSK signal waveforms.

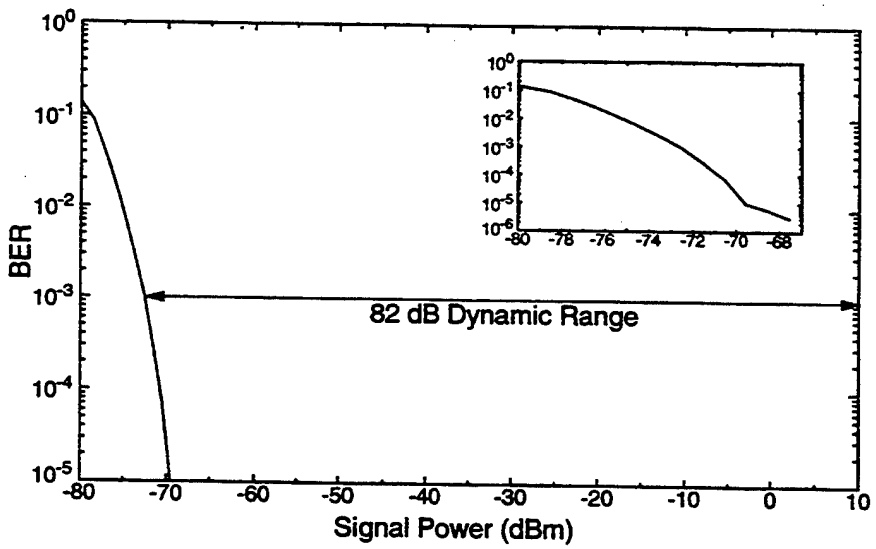


Fig. 14. Measured dynamic range for the binary FSK detector.

it is not utilized. The scaler, rounding device, and loop filter operate in the same manner as those in the time tracking loop.

D. I/O Port Clocking Scheme and Testability

A gated clock is produced to drive those registers which are not accessed frequently, such as the input and output buffer registers and those registers which store the programmable coefficients. Compared with the circuit shown in Fig. 10(a), this gated clock [Fig. 10(b)] will only charge the large gate and routing capacitances of the registers when new data is loaded. On the other hand, if the master clock is used directly to drive these registers, this switching power dissipation will always exist. HSPICE simulations showed that about 10% of the total chip power dissipation is saved by using this gated clock scheme. The skew between the off-chip and on-chip clocks will not affect the operation of these registers. The data bus and the address bus (which produces the LOAD signal) can also share the same set of input pads in a time-multiplexed manner [Fig. 10(c)]. The testability of the loops is made possible by using readable and writable registers in the loop filters and control blocks.

IV. Baseband Signal Processing Blocks Implementation

The baseband transmitter and all of baseband signal processing with the exception of the channel codec and the DDFS/DAC are implemented in two XILINX FPGA chips for the prototype handset to demonstrate a point-to-point communication link in the 902–928 MHz ISM band. This choice is mainly due to the ease in design modifications provided by FPGA's. For the multirate processing and hardware framing blocks of the baseband signal processor, timing control is extremely critical. Timing problems are carefully avoided by inserting extra delays when necessary in the FPGA chips. One of the chips contains the entire controller block including the 16-b frame counter and most of the receive path. The other FPGA implements the transmit path and the deinterleaver. This partitioning gives a rather equal utilization of the resources in each chip and minimizes the interconnections between the two chips.

V. Results

The functionality of the baseband receiver chip was verified with the Tektronix LV500 ASIC tester. It was found that 96% of the 25 chips received from the MOSIS fabrication run passed the functionality test. The minimum operating voltages for these functional chips were also measured and shown in a histogram (Fig. 11) for a 80 kbaud data rate and a 88 times oversampling rate. The measured power dissipations for different supply voltages demonstrate a clear quadratic power reduction achieved by voltage scaling (Fig. 12). About 58% of the power consumption can be saved by using the minimum operating voltage of 1.95 v instead of 3 v while still meeting the desired circuit speed.

The dynamic range of a 1-b binary FSK detector with $F_{\text{baud}} = 160$ kHz and an oversampling ratio (N) of 64 was measured. In the test setup, a quadrature DDFS/DAC combination is used to convert the test data sequence generated by the transmitter of a bit error rate (BER) tester into the baseband binary FSK tones, $F_{\text{tone}} = 160$ kHz (Fig. 13). A pair of limiting amplifiers is then used to hard-limit the quadrature analog FSK tones for 1-b signal correlation. The hard-decision output from the demodulator chip is sent to the receiver of the BER tester for measurements. In order to vary the dynamic range of the input signal, power attenuators were inserted between the DAC's and the hard limiters. For a bit error rate of 10^{-3}, the measured dynamic range is 82 dB (Fig. 14), which is sufficient for the radio channels encountered in most wireless applications.

Both of the synchronization loops were tested using the self-test mode of the baseband transceiver system. The performance of the time tracking loop (TTL) was verified in a closed loop fashion while receiving a pseudorandom data sequence. The output of the phase detector (PD) was captured immediately after the TTL was enabled. The trajectory of the PD output is shown in Fig. 15 and demonstrates the operation of the TTL as expected. The FD output of the frequency tracking loop in the baseband receiver chip was also measured by inserting known frequency offsets between transmitter and receiver. These frequency offsets were implemented by modifying the frequency control word for the transmitter DDFS. The corresponding FD outputs were collected by a logic analyzer and averaged out to characterize the function of the FD. The measured S-curve of the FD shown in Fig. 16 covers a wide range of frequency offsets up to ±50 kHz, which is more than half the baud frequency.

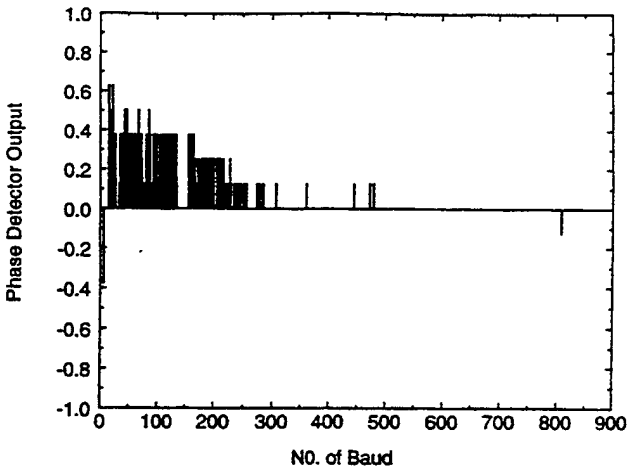


Fig. 15. Measured trajectory of the phase detector output of TTL.

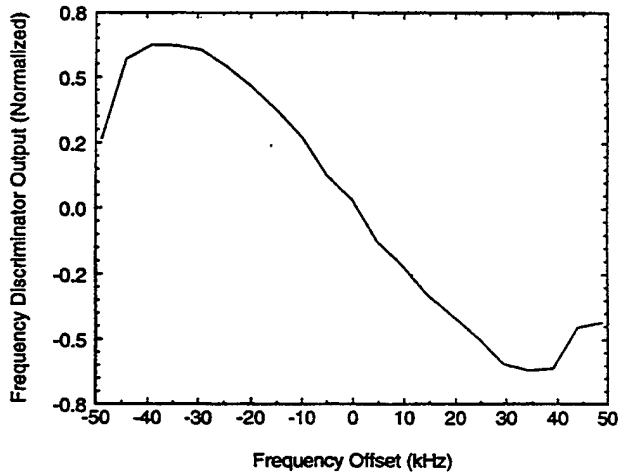


Fig. 16. Measured S-curve of the frequency detector output of FTL.

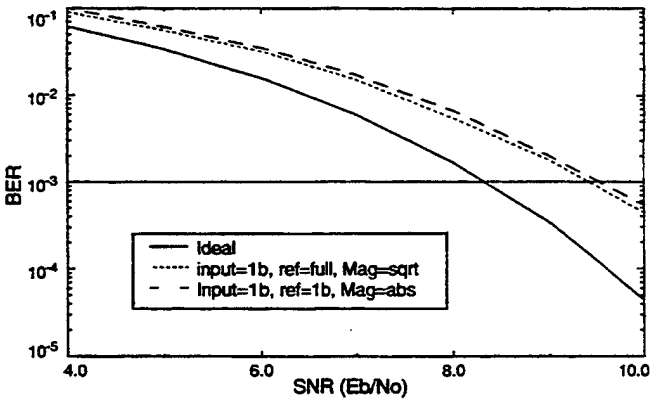


Fig. 17. Implementation loss of the 1-b correlation FSK detector, 4-FSK in AWGN.

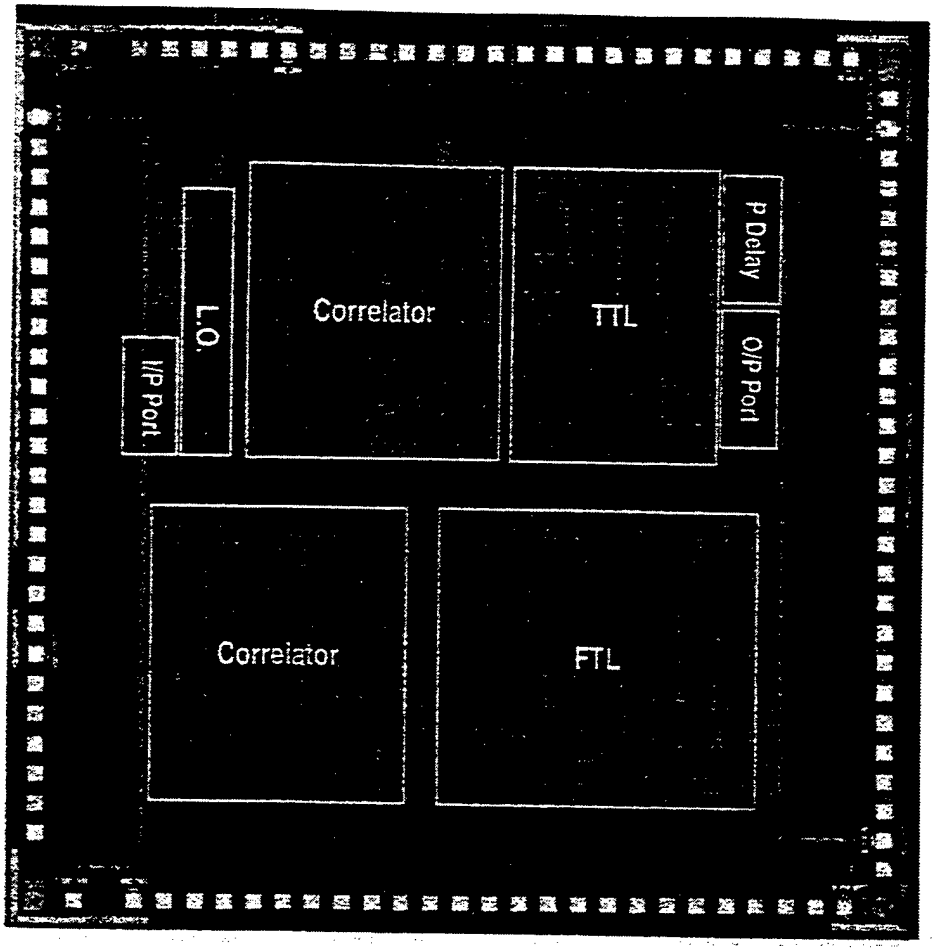


Fig. 18. Baseband receiver chip microphotograph.

TABLE I
KEY FEATURES OF THE BASEBAND RECEIVER CHIP

Technology	1 μm CMOS (HPCMOS34)
Clock Frequency	10 MHz
Date Rates	2 - 160 kbps
Modulation Formats	Binary/quaternary noncoherent FSK
Antenna Diversity Combining	Dual branch, equal-gain combining
Maximum Hop Rate	160 khop/s
Oversampling Rate (Fclk/Fbaud)	2 - 160
Loop Filter Coefficients	Programmable, powers of 2
Power Dissipation @ 3V	4.5 mW @ 10 MHz
Transistor Count	48,000
Core area	3.9 mm x 3.9 mm
Die area	5.4 mm x 5.4 mm

The implementation loss of our multiplierless FSK detector was also studied via bit-level simulations. Simulation results show about 1.2 dB of implementation loss for the quaternary FSK case in an additive white Gaussian noise (AWGN) channel from the ideal full-precision correlation detector. Given that the input signals are already hard-limited, extra degradation due to 1-b signal correlation or absolute-value addition is minimal (only about 0.1 dB more), as shown in Fig. 17. This result justifies our baseband multiplierless FSK correlation detector for direct-conversion receivers. The key features of this baseband receiver chip are listed in Table I and its chip photo is shown in Fig. 18.

VI. CONCLUSIONS

A low-power all-digital optimal correlation detector IC for FSK has been presented. A low-power design is achieved by using both architecture-level optimizations and circuit-level techniques. At the architecture level, expensive (both in power and area) digital multipliers are replaced by simpler circuits such as 1-b correlators and absolute-value summing nodes. The use of a 1-b-output NCO, rather than high-precision DDFS's, as the FSK reference tone generator, also contributes to a low-power design. At the circuit level, parallelism for high frequency blocks combined with logic simplifications is employed to further decrease the power consumption. A robust gated clocking scheme used in the I/O ports and peripheral circuits minimizes the capacitive loading of the high frequency clock signal, thereby reducing the overall power consumption to only 4.5 mW at a 10 MHz clock rate. For the FSK detector, only the front-end correlators are high-frequency components. Thus, time-multiplexing of low frequency blocks is exploited to reduce the overall chip area. This results in a core area of

only 15.2 mm^2 for the digital detector. The rest of the baseband signal processing is implemented in two XILINX XC4010 chips. The proposed baseband transceiver is well-suited for use in low power FH/SS CDMA handheld communications applications.

ACKNOWLEDGMENT

The authors wish to thank all members of the FH/SS project team at UCLA for their technical contributions.

REFERENCES

[1] J. Min, A. Rofougaran, V. Lin, M. Jensen, H. Samueli, A. A. Abidi, G. Pottie, and Y. Rahmat-Samii, "A low power handheld frequency-hopped spread spectrum transceiver hardware architecture," in *Wireless Personal Communications: Trends and Challenges*, T. S. Rappaport, B. D. Werner, and J. H. Reed, Eds. Norwell, MA: Kluwer Academic, pp. 119–126, 1994.

[2] J. Min, A. Rofougaran, H. Samueli, A. A. Abidi, "An all-CMOS architecture for a low-power frequency hopped 900 MHz spread spectrum transceiver," in *Proc. of IEEE Custom Integrated Circuits Conf.*, May 1994, pp. 379–382.

[3] G. Chang, A. Rofougaran, M.-K. Ku, A. A. Abidi, and H. Samueli, "A low-power CMOS digitally synthesized 0–13 MHz sinewave generator," in *Tech. Dig. IEEE 1994 Int. Solid-State Circuits Conf.*, Feb. 1994, pp. 32–33.

[4] J. F. Wilson, R. Youell, T. H. Richards, G. Luff, and R. Pilaski, "A single-chip VHF and UHF receiver for radio paging," *IEEE J. Solid-State Circuits*, vol. 26, no. 12, pp. 1944–1950, 1991.

[5] G. Schultes, A. L. Scholtz, E. Bonek, P. Veith, "A new incoherent direct conversion receiver," in *Proc. 40th IEEE Vehicular Technology Conf.*, May 1990, pp. 668–674.

[6] P. Kreuzgruber, "A class of binary FSK direct conversion receivers," in *Proc. 44th IEEE Vehicular Technology Conf.*, June 1994, pp. 457–461.

[7] J. G. Proakis, *Digital Communications*, 2nd ed. New York: McGraw-Hill, 1989.

[8] J. Min, H-C. Liu, A. Rofougaran, S. Khorram, H. Samueli, and A. A. Abidi, "Low power correlation detector for binary FSK direct-conversion receivers," *Electron. Lett.*, vol. 31, pp. 1030–1032, June 22, 1995.

[9] F. Lu, H. Samueli, J. Yuan, and C. Svensson, "A 700-MHz 24-bit pipelined accumulator in 1.2-μm CMOS for application as a numerically controlled oscillator," *IEEE J. Solid-State Circuits*, vol. 28, no. 8, pp. 878–886, Aug. 1993.

[10] F. M. Gardner, *Phaselock Techniques*, 2nd ed. New York: Wiley, 1979.

[11] S. Khorram, A. Rofougaran, and A. A. Abidi, "A CMOS limiting amplifier and signal-strength indicator," presented at *Symp. VLSI Circuits*, Kyoto, Japan, June 1995.

[12] R. H. Barker, "Group synchronizing of binary digital systems," in *Communication Theory*. London: Butterworth, 1953, pp. 273–287.

[13] V. S. Lin and G. J. Pottie, "Channel coding for a frequency-hopped wireless transceiver," in *Proc. 1993 Canadian Workshop on Information Theory*, June 1993, pp. 109–124.

[14] H-C. Liu, J. Min, and H. Samueli, "A low-power baseband receiver IC for frequency-hopped spread spectrum applications," in *Proc. IEEE Custom Integrated Circuits Conf.*, Santa Clara, CA, May 1995, pp. 311–314.

[15] A. P. Chandrakasan, S. Sheng, and R. W. Brodersen, "Low-power CMOS digital design," *IEEE J. Solid-State Circuits*, vol. 27, no. 4, pp. 473–483, Apr. 1992.

Huan-Chang Liu (S'91) was born in Taipei, Taiwan, R.O.C., on March 20, 1966. He received the B.S. degree in electrical engineering from the National Taiwan University in 1988, the M.S. degree in electrical engineering, and the Engineer degree from the University of California, Los Angeles, in 1992 and 1995, respectively.

From 1988 to 1990 he served as an Electronics Instructor in the Chinese Navy Communications and Electronics School. Since 1991, he has been a research assistant at UCLA. He is currently pursuing the Ph.D. degree at UCLA. His research interests include VLSI implementation of digital signal processing algorithms, low power circuit techniques, and digital communication systems.

Jonathan S. Min received the B.S. and M.S. degrees in electrical engineering and computer science from the University of California, Berkeley, in 1987 and 1989, respectively, and the Ph.D. degree in electrical engineering from the University of California, Los Angeles, in 1995.

From 1989 to 1991, he was with Rockwell International, Anaheim, CA, where he was involved in the development of space electronic systems such as a multichannel GPS receiver and satellite transceivers. Currently, he is with Broadcom Corporation, Irvine, CA, working on the design and development of high-speed cable modems. His research interest are in the areas of wireless/wireline communications systems design, spread spectrum transceivers, digital signal processing, and VLSI architecture for realizing advanced digital communications systems.

Henry Samueli (S'75–M'81) was born in Buffalo, NY, on September 20, 1954. He received the B.S., M.S., and Ph.D. degrees in electrical engineering from the University of California, Los Angeles, in 1975, 1976, and 1980, respectively.

From 1980 to 1985 he was with TRW, Inc., Redondo Beach, CA where he was a Section Manager in the Digital Processing Laboratory of the Electronics and Technology Division. His group was involved in the hardware design and development of military satellite and digital radio communication systems. From 1980 to 1985 he was also a part-time Instructor in the Electrical Engineering Department at UCLA. In 1985, he joined UCLA full-time where he is currently a Professor in the Electrical Engineering Department. His research interests are in the areas of digital signal processing, communications systems engineering, and CMOS integrated circuit design for applications in high-speed data transmission systems. In 1988, he co-founded PairGain Technologies, Tustin, CA, a telecommunications equipment manufacturer, and in 1991, he co-founded Broadcom Corporation, Irvine, CA, a communications IC supplier.

Dr. Samueli is the recipient of the 1988/1989 TRW Excellence in Teaching Award of the UCLA School of Engineering and Applied Science and the Meritorious Paper Award of the 1991 Government Microcircuit Applications Conference.

An 800-MHz Quadrature Digital Synthesizer with ECL-Compatible Output Drivers in 0.8 mm CMOS

L.K. Tan, E. Roth, G. Yee, and H. Samueli

IEEE Journal of Solid-State Circuits
vol. 30, no. 12, pp. 1463-1473
December 1995

An 800-MHz Quadrature Digital Synthesizer with ECL-Compatible Output Drivers in 0.8 μm CMOS

Loke Kun Tan, Edward W. Roth, Gordon E. Yee and Henry Samueli, *Member, IEEE*

Abstract— **An 800 MHz quadrature direct digital frequency synthesizer (QDDFS4) chip is presented. The chip synthesizes 12 b sine and cosine waveforms with a spectral purity of −84.3 dBc. The frequency resolution is 0.188 Hz with a corresponding switching speed of 5 ns and a tuning latency of 47 clock cycles. The chip is also capable of frequency and phase modulation. ECL-compatible output drivers are provided to facilitate I/O compatibility with other high speed devices. A high gain amplifier at the clock input enables the QDDFS4 chip to be clocked with ac-coupled RF signal sources with peak-to-peak voltage swings as small as 0.5 V. The 0.8 μm triple level metal N well CMOS chip has a complexity of 94 000 transistors with a core area of 5.9 × 6.7 mm^2. Power dissipation is 3 W at 800 MHz and 5 V.**

I. Introduction

TRADITIONAL designs of high bandwidth frequency synthesizers employ the use of a phase-locked-loop (PLL). However, the presence of a feedback loop in a PLL implementation significantly increases the acquisition time hence impacting its ability to simultaneously provide fast frequency switching and high frequency resolution [1]. Direct digital frequency synthesizers (DDFS's) do not exhibit these disadvantages since their design is purely feedforward. Thus, the switching speed of a DDFS is essentially instantaneous. Other advantages inherent in the DDFS approach include its ability to provide fine frequency resolution and phase coherent frequency switching. Furthermore, the amount of phase noise in a DDFS is primarily limited by that of the reference clock source. The primary drawback to the DDFS approach is the requirement for high-speed, high resolution D/A converters. In times past, DDFS's have been considered a low speed method of frequency synthesis since their performance is directly limited by the availability of fast digital integrated circuit (IC) processes. Fortunately, recent advances in IC technology have brought about considerable progress in this area. Advanced GaAs and Si-bipolar IC processes have resulted in DDFS designs with very high operating frequencies [2], [3]. However, the esoteric nature as well as the lower yields of these processes result in relatively costly solutions. This paper presents a high performance 800-MHz monolithic quadrature DDFS design that achieves a spectral purity of −84.3 dBc using a generic 0.8-μm CMOS process. The high performance of the QDDFS4 was achieved through the use of a parallel architecture.

Manuscript received May 22, 1995; revised July 25, 1995. This work was supported in part by research grants from the University of California MICRO Program, Hewlett-Packard, TRW Electronic Systems Group, Hughes Space and Communications Group, and the Advanced Research Projects Agency.

L. K. Tan is with Broadcom Corporation, Irvine, CA 92718 USA.

E. W. Roth, G. E. Yee, and H. Samueli are with Integrated Circuits and Systems Laboratory, University of California, Los Angeles, CA 90024 USA.

IEEE Log Number 9416003.

II. Chip Architecture

A. Overall Architecture

The basic architecture of the QDDFS4 chip consists of an overflowing L-bit accumulator (or phase accumulator) to generate the phase argument of the sine function generator. Each overflow of the phase accumulator represents one period of a sine wave. The input word (Frequency Control Word) to the phase accumulator controls the frequency of the generated sine waveform. The sine function generator is a ROM lookup table that stores the sine samples. The use of parallelism to attain high throughput is common in VLSI design and has been utilized in [4] and [5] for DDFS applications. Our chip has been designed with four parallel ROM lookup tables to achieve four times the throughput of a single DDFS. The size of the ROM table is $2^8 \times$ 12 b. Each ROM table generates every fourth sample of a sine waveform. This is performed by applying four times the Frequency Control Word (FCW) to each ROM table and also adding a phase offset equal to 0, 1, 2, and 3 times the Frequency Control Word to each corresponding ROM table. Last, a 4-to-1 MUX is used to alternately select the outputs of the four ROM's to reconstruct the final sine waveform. Thus, the only block that operates at the maximum clock rate of the chip is the 4-to-1 MUX. All other blocks operate at a quarter of the reference clock frequency. The detailed architecture of this design is illustrated in Fig. 1. A chip that uses only one ROM table has been fabricated and tested to perform at 200 MHz [6]. Using the parallel architecture with four ROM tables, the QDDFS4 chip attains the targeted speed of 800 MHz.

The QDDFS4 has a phase accumulator wordlength of 32 b. When clocked at 800 MHz, a minimum frequency resolution of 0.188 Hz is attained. The 32 b phase accumulator output is truncated to 14 b prior to addressing the ROM lookup table, resulting in a −84.3 dBc worst case spurious rejection based on the fact that phase accumulator truncation is the dominant source of spurious noise [7].

B. Quadrature Outputs

The QDDFS4 chip takes advantage of the quarter wave symmetry of a sine wave to reduce ROM storage requirements. Thus, only sine samples from 0 to $\pi/2$ are stored. To produce

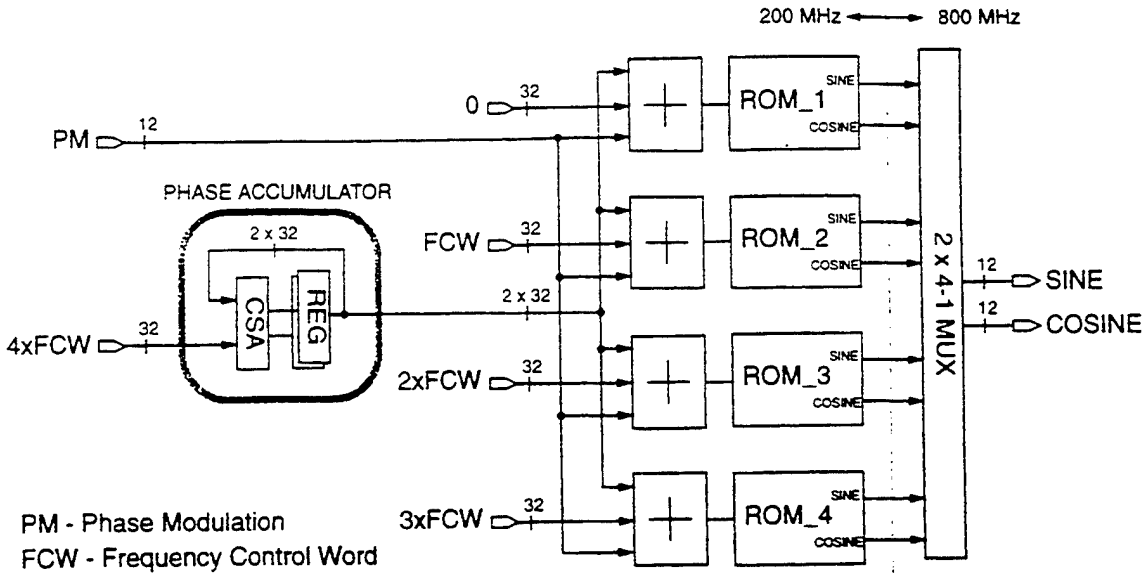


Fig. 1. Detailed top level architecture of QDDFS4.

quadrature outputs, rather than having a separate ROM lookup table for the cosine waveform, this design takes advantage of the eighth wave symmetry of the sine and cosine functions. Since sine samples from 0 to $\pi/4$ are identical to cosine samples from $\pi/4$ to $\pi/2$, and cosine samples from 0 to $\pi/4$ are identical to sine samples from $\pi/4$ to $\pi/2$ thus, one ROM lookup table containing only sine samples has sufficient information to produce both sine and cosine waveforms. Thus, the only additional hardware cost of generating quadrature outputs is the use of a 2-to-1 MUX to select the appropriate ROM table contents to produce the sine and cosine samples [6].

C. Modulation Formats

This chip has modulation capabilities that include frequency modulation and phase modulation. Frequency modulation is performed by directly modulating the Frequency Control Word, thus, no additional hardware is needed to implement this feature. Phase modulation is accomplished by adding a phase offset to the phase accumulator output before addressing the ROM lookup table. In hardware, this amounts to incorporating an extra addition stage. The QDDFS4 accepts a 12 b word for phase modulation. Fig. 1 illustrates these modulation capabilities. To reduce input pad count, the 12 b phase modulation word is multiplexed with the 12 MSB's of the Frequency Control Word.

III. Circuit Design

The size of the ROM lookup table is $2^8 \times$ 12 b. The decode logic of the ROM was designed using pseudo-NMOS logic while the ROM basic cell uses dynamic Domino logic. The carry propagate adder that produces the final phase word to address the ROM lookup table uses a conditional sum architecture. Details pertaining to the design of the conditional sum adder and the ROM lookup table are described in [6].

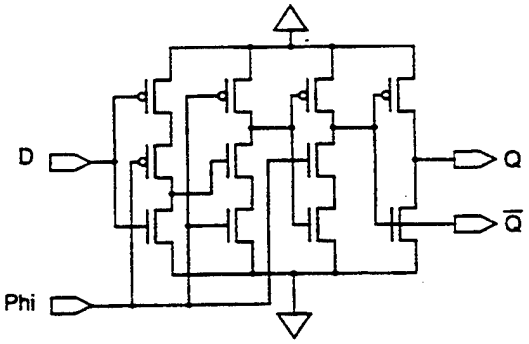


Fig. 2. Single-phase positive-edge triggered register.

A. Register Design

The QDDFS4 chip uses only dynamic positive-edge triggered single-phase registers. A schematic of this register is illustrated in Fig. 2. Standard two-phase clocking schemes for high speed operation are difficult to design since small amounts of clock skew directly limit speed. Single-phase registers do not exhibit a sensitivity to clock skews and are thus able to achieve higher clock frequencies [8]. SPICE simulations of this register were performed over variations in transistor models, temperature, and power supply levels to compute T_s and $T_{\text{clk}-Q}$, where T_s is the register setup-time and $T_{\text{clk}-Q}$ is the register clock-to-Q delay. A plot of $T_s + T_{\text{clk}-Q}$ versus T_s is illustrated in Fig. 3. Each line in the plot describes a particular simulation condition with a fixed transistor model, fixed temperature, and fixed power supply voltage. The entire collection of simulation cases characterizes the spread of timing information associated with this register. The data obtained from this plot is summarized in Table I. From this table, it can be seen that the worst case setup time is $T_s = 0.13$ ns. If the setup-time T_s were reduced below this value, at least one simulation case fails to latch data correctly. Using $T_s = 0.13$ ns as a reference point, $T_{\text{clk}-Q} + T_s$ varies

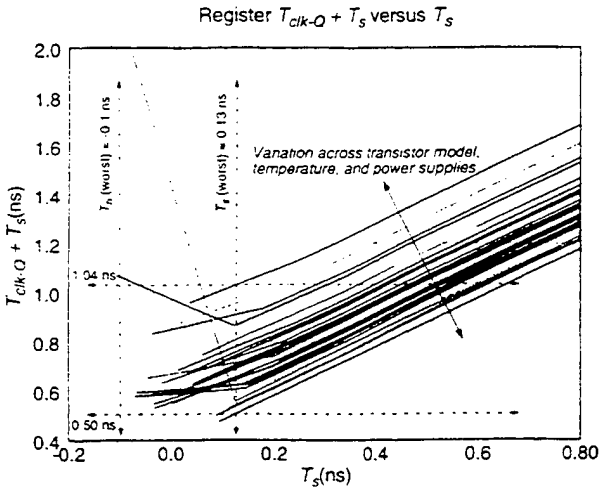


Fig. 3. SPICE simulation of single-phase positive-edge triggered register.

TABLE I
SUMMARY OF REGISTER TIMING PARAMETERS BASED ON SPICE SIMULATIONS

Timing Parameter	Fast Case (ns)	Nominal Case (ns)	Slow Case (ns)
$T_{clk-Q} + T_s$ (@ T_s = 0.13ns)	0.50	0.77	1.04
T_{clk-Q} (linear region)	0.37	0.64	0.91
Timing Parameter	Worst Case (ns)		
T_h	0.10		
T_s	0.13		

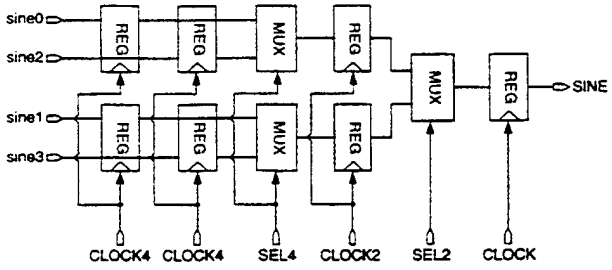


Fig. 4. Block diagram of 4-to-1 MUX.

from 0.50–1.04 ns across all simulation conditions. As long as the register setup time is not violated, the relationship between $T_{\text{clk}-\text{Q}} + T_s$ and T_s is linear. Hence, $T_{\text{clk}-\text{Q}}$ can be calculated for this region of operation. Subtracting $T_s = 0.13$ ns from the first row of Table I results in $T_{\text{clk}-\text{Q}}$, which ranges from 0.37–0.91 ns. The other important timing parameter in this simulation is the worst case hold time $T_h = 0.10$ ns. If this hold time is violated, at least one simulation case will exhibit data flash-through.

B. 4-to-1 MUX Design

The design of the 4-to-1 MUX is challenging since it is the only block operating at the 800 MHz clock rate. The circuit schematic of this 4-to-1 MUX is shown in Fig. 4. Two such MUXes are used; one for generating sine waveforms and one for cosine waveforms. The signals Sine0, Sine1, Sine2, and Sine3 are the four staggered sine samples used

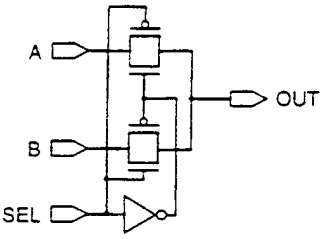


Fig. 5. Transistor schematic of 2-to-1 MUX.

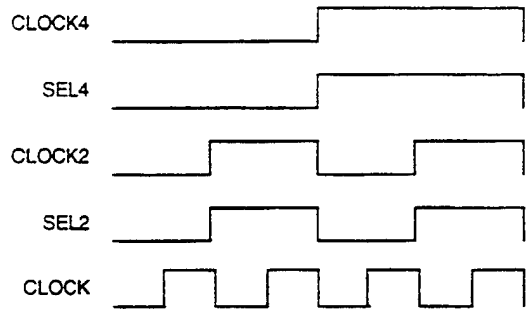


Fig. 6. Timing of clock and select signals.

to reconstruct the final sine waveform. The 4-to-1 MUX consists of a cascade of two 2-to-1 MUXes. Pipeline stages are placed in between these two stages for maximum throughput. All pipeline registers are single-phase positive-edge triggered registers. A transistor schematic of the 2-to-1 MUX is shown in Fig. 5. The clock signals CLOCK, CLOCK2, and CLOCK4 are the input clock, input clock divided by 2, and input clock divided by 4, respectively. The timing relationship of these clocks and select signals are shown in Fig. 6.

C. Clock Generation

The timing of the clock and select lines is extremely critical. The circuit used for the generation of these signals is illustrated in Fig. 7. A divide-by-2 flip-flop is used twice to generate the CLOCK2 and CLOCK4 signals. These two circuits are cascaded to create an asynchronous divide-by-4 circuit for the generation of the CLOCK4 signal. The use of an on-chip high gain amplifier enables this chip to be clocked with RF signal sources that are capable of generating small voltage swing ac-coupled signals in the GHz frequency range. This circuit accepts voltage swings as small as 0.5 $V_{\text{p}-\text{p}}$ and also exhibits a low sensitivity to noise [9]. The inset in Fig. 7 illustrates the transistor schematic of this design. The capability of comparing the threshold of the clock signal to a dc input reference voltage V_{ref} gives this design an insensitivity to common-mode threshold shifts. The output of the clock amplifier labeled N2 drives a CMOS inverter that further increases the gain of the combination. The manner in which this clock amplifier is connected to the clock generation circuit of the chip is illustrated in Fig. 7. Back-annotated SPICE simulations predicted that the clock generation circuit is capable of producing on chip CMOS clocks in excess of 800 MHz. This result is illustrated in Fig. 8 where the top trace is the 0.5 $V_{\text{p}-\text{p}}$ input reference clock. The final

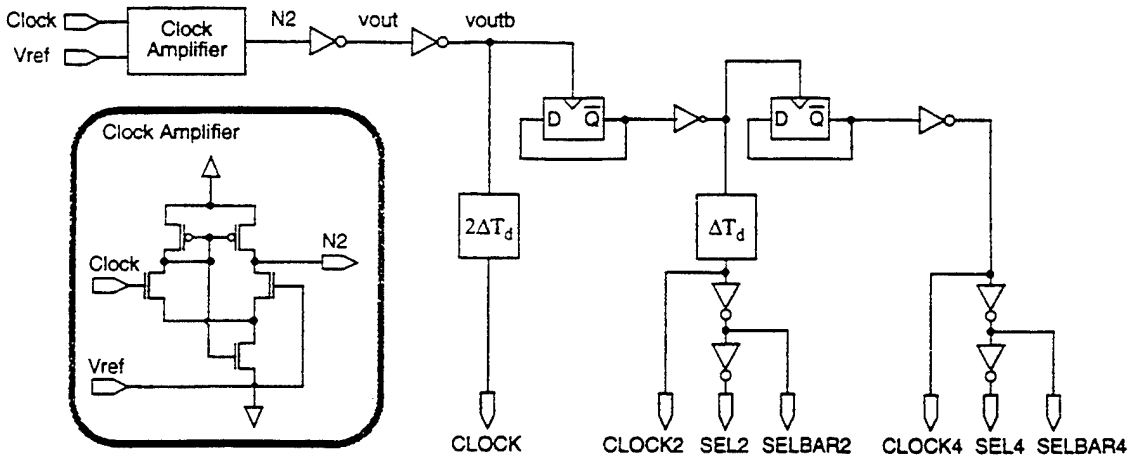


Fig. 7. Block diagram of clock generation circuit.

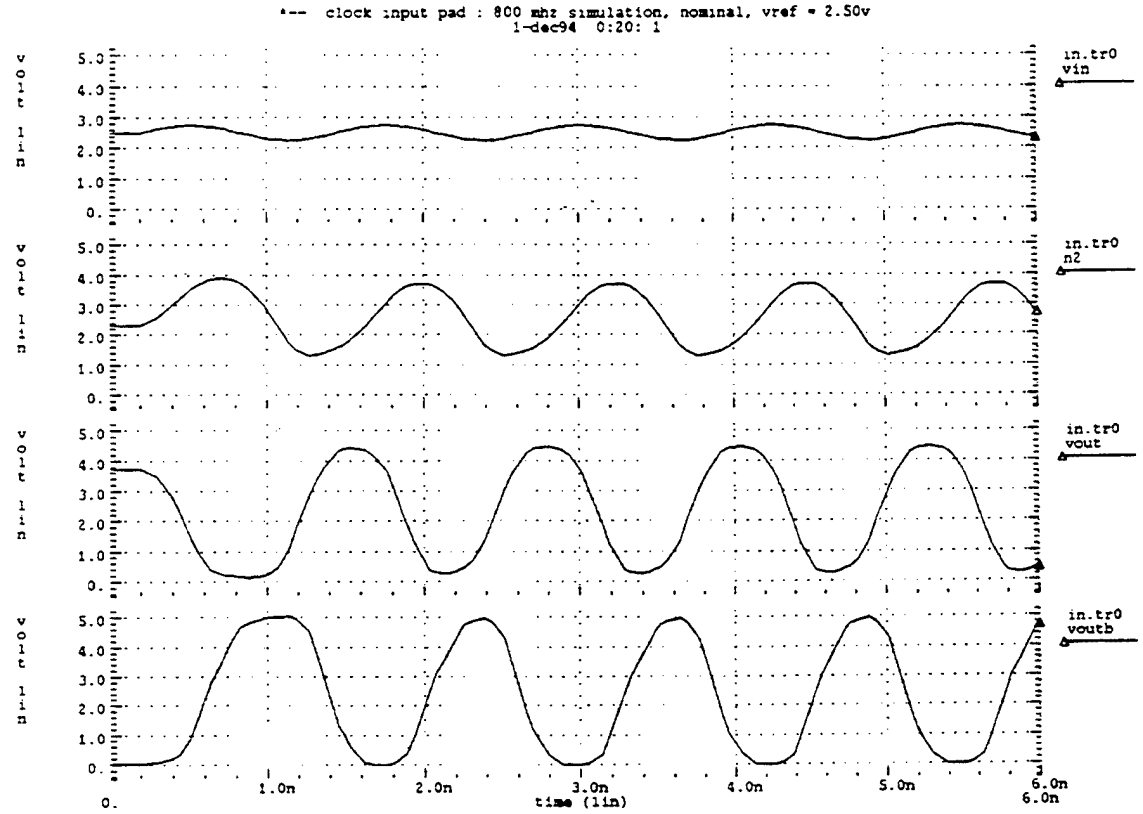


Fig. 8. SPICE simulation of clock amplifier at 800 MHz.

waveform Voutb has full 0–5 V signal swing and is thus fully CMOS compatible. The simulation was performed using nominal transistor models, at 65°C, 5 V power supplies, and worst case extracted parasitics.

The use of an asynchronous divide-by-4 circuit results in register clock-to-Q delays between the three clock signals CLOCK, CLOCK2, and CLOCK4. A synchronous design would have been a safer choice where minimizing clock skews is concerned; however, such a design could never achieve the necessary 800 MHz throughput. The divide-by-2 circuit is essentially a dynamic positive-edge triggered register of Fig. 2 with the $\bar{Q}$ output feeding back to the D input. This circuit was simulated across models, temperature, and power supply voltages to ensure robust operation. The blocks labeled ΔT_d and $2\Delta T_d$ are used to compensate for the asynchronous divider delays. A simple timing analysis indicates that it is possible to increase the throughput of the 4-to-1 MUX by over-compensating these delays. This timing analysis can be

described using a block diagram showing the final stage of the 4-to-1 MUX and its associated timing diagram. This is illustrated in Fig. 9. The timing delays in this figure assume that the delay through a 2-to-1 MUX is negligible. This is a valid assumption since the critical path through this block is an inverter driving the drain-source region of a transmission gate that is connected to a purely capacitive load on its other side. Hence, the delay through the 2-to-1 MUX is very small. From Fig. 9, two critical timing parameters T_{sb} and T_{hb} must be met in order for the circuit to function properly. T_{sb} is the remaining amount of setup time, and T_{hb} is the hold time available when data is latched from node B. These two parameters must be smaller than the actual required setup (T_s) and hold (T_h) times of a register. From Fig. 9, the timing relations governing T_{hb} and T_{sb} are

$$T_{clk-Q} - T_d + 2T_{inv} = T_{hb} \geq T_h \quad (1)$$

$$T_p - (T_{clk-Q} - T_d + T_{clk-Q}) = T_{sb} \geq T_s \quad (2)$$

where T_{clk-Q} is the clock-to-Q delay of a register, and T_{inv} is the delay of an inverter gate, and T_d is the delay inserted between CLOCK and CLOCK2 to compensate for the clock-to-Q delay that exists between these two signals. T_{inv} has been simulated to be 0.058 ns, 0.088 ns, and 0.165 ns under fast, nominal, and slow conditions, respectively. Table II summarizes T_{inv} and T_{clk-Q} for these cases. The task here is to determine an appropriate T_d measured in number of inverter delays such that T_p is minimized and furthermore, both (1) and (2) are satisfied. By substituting T_h, T_s, T_{inv} and T_{clk-Q} from Tables I and II into both (1) and (2), one can determine a range of T_d such that setup and hold times are satisfied. Plots of frequency ($1/T_p$) versus T_d (normalized to number of inverter delays) for slow, nominal, and fast conditions are illustrated in Fig. 10. The shaded sections in these three plots indicate acceptable regions of operation where both (1) and (2) are not violated. Choosing an operating point further to the right or equivalently using a larger T_d results in increasing maximum throughput of the 4-to-1 MUX. The limit of this choice is the right boundary of operation, which indicates that a hold time requirement is violated. To accommodate process variations as well as varying operating conditions, it is important that the value of T_d be chosen so as to satisfy all three plots. Furthermore, T_d has to be an even number of inverter delays; otherwise, the clock signal will be inverted. Thus, based on these results, T_d is chosen to be six inverter delays. Also based on the results of Fig. 10, the maximum frequency of operation f_{max} ranges from 880 MHz–1.9 GHz depending on process variations and operating conditions. It is doubtful that a throughput of 1.9 GHz can be achieved since other factors not accounted for in the timing model might prove to be dominant under this condition.

D. Output Driver Design

Due to the chip's very high operating frequency, the design of the output drivers is crucial. Since the large switching currents of CMOS drivers generate substantial di/dt noise, ECL compatible drivers are not only desirable, but also

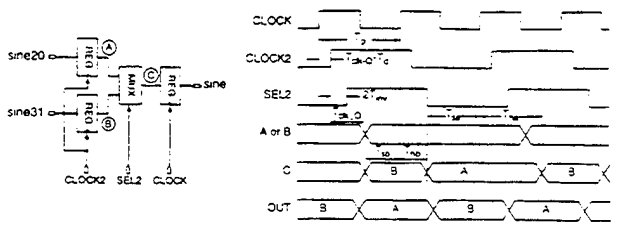


Fig. 9. 4-to-1 MUX final stage block diagram and associated timing diagram.

TABLE II
SUMMARY OF SIMULATION RESULTS OF T_{inv} AND T_{clk-Q}

Simulation Condition	T_{inv} (ns)	T_{clk-Q} (ns)
Fast, 0°C, 5.5V	0.058	0.37
Nominal, 25°C, 5.0V	0.088	0.64
Slow, 125°C, 4.5V	0.165	0.91

necessary. ECL drivers also facilitate direct I/O compatibility with commercial high-speed D/A converters.

The ECL compatible output driver (Fig. 11) consists of output buffers and a control unit similar to the design described in [10]. A single pull-up transistor at the output stage of the buffer cannot be used because it may not satisfy the small tolerances of an ECL level output (typically within 0.08 V) due to process variations. Fig. 12 shows the detailed circuit of the ECL output driver where a current source is used as the output stage. When V_{in} is *low*, the current mirror ($M6$ and $M7$) is off, and V_{out} is simply $V_{L(ECL)}$ (~ -1.7). When V_{in} is *high*, the current source is switched on, thus causing a constant voltage across a matching resistor R_L. A second current mirror ($M1$ and $M2$) is used to decouple the switched current mirror from I_{ref}, the input reference current generated by the control unit.

One of the improvements made to this circuit was the addition of M3. This PMOS transistor acts as a current steering device to dampen the ground bounce on the power supply. $M3$ and $M7$ are complementary relative to each other, that is, if one is on, the other is off, and vice versa. As a result, the output buffer draws finite current in the *low* state. This effectively reduces the difference in current drawn between the *high* and *low* states of the output buffer and thus reduces inductive ground bounce. Simulations have indicated that the addition of $M3$ can reduce inductive ground bounce by as much as 1 V.

Unacceptable variations in the *high* output level may be caused by the degeneration of the current mirror due to channel-length modulation. Therefore, a feedback control unit is used to compensate for technology parameter variations. The control unit consists of the transistor schematic of Fig. 13 as well as a dummy output driver (Fig. 11) whose input is fixed at Vdd to generate a permanent *high* output level (Out[0]). That *high* output level is compared with the desired level $V_{H(ECL)}$, and the result is used to control the reference current source (I_{ref0}) at its output regardless of current mirror imperfections. Since the other output drivers receive the same reference current ($I_{ref2} = I_{ref1} = I_{ref0}$), their *high* output level is also controlled by $V_{H(ECL)}$.

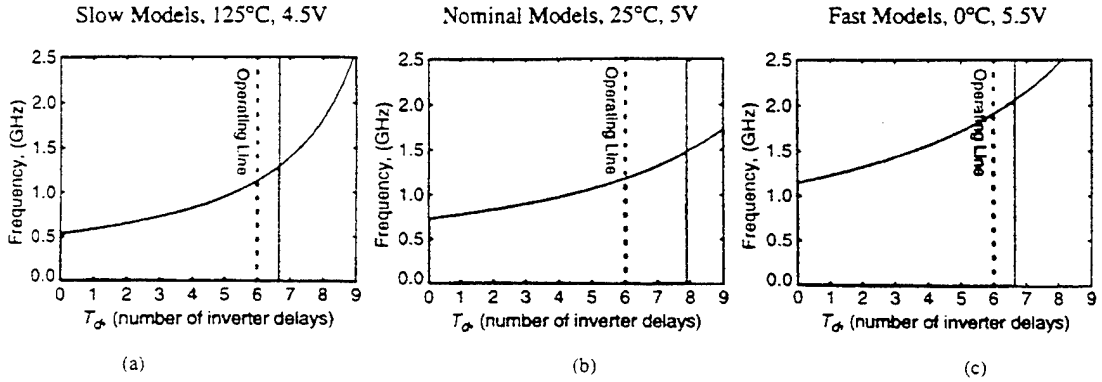


Fig. 10. SPICE simulation summary of 4-to-1 MUX region of operation.

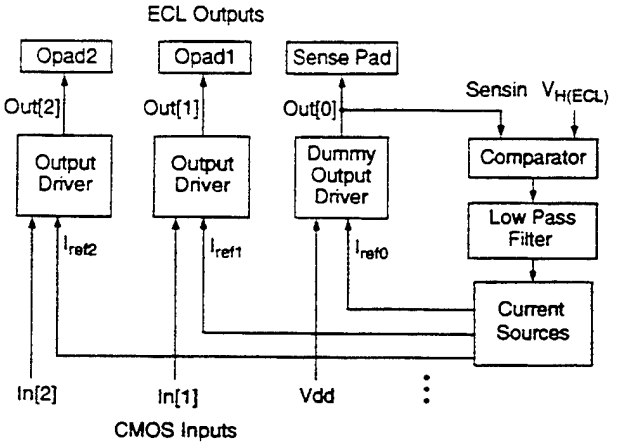


Fig. 11. Block diagram of CMOS-to-ECL level output drivers.

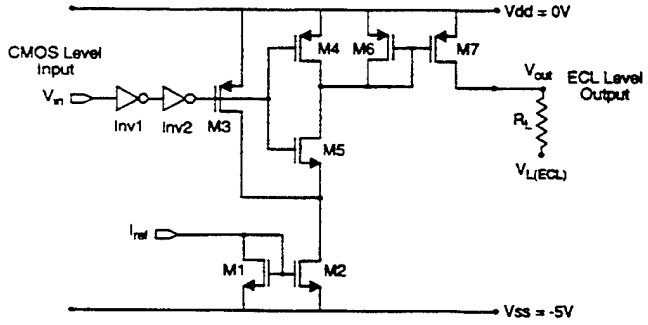


Fig. 12. Transistor schematic of output buffer.

The configuration of the control circuit was modified to ensure that the comparator of Fig. 13 possesses only one dominant pole located at node Z. Additional current sources will only move that pole down in frequency and further away from any secondary pole. Thus, system stability is maintained while keeping the number of desired current sources flexible. Moreover, the connection of a capacitor from node Z to V_{dd} helps maintain a constant gate-source voltage V_{gs} of the PMOS current sources and thus a constant I_{ref} for the output drivers. This capacitor has no effect on the speed of the driver since the control unit itself is not operating at switching frequencies.

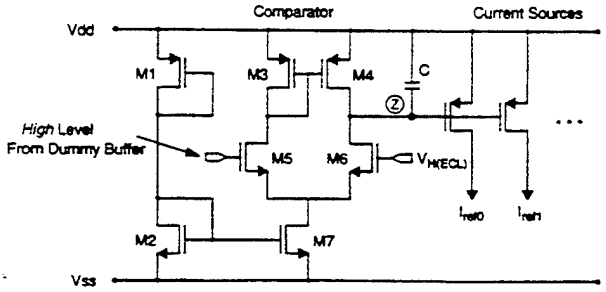


Fig. 13. Transistor schematic of control unit for ECL output drivers.

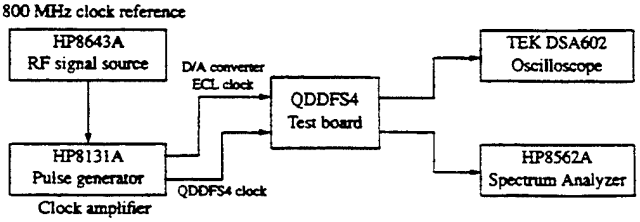


Fig. 14. QDDFS4 test setup.

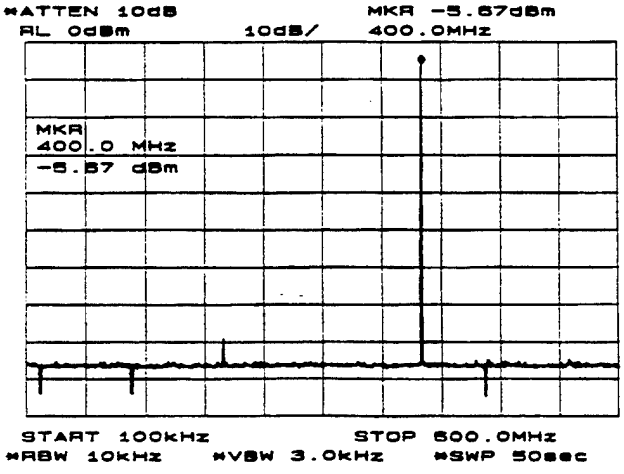


Fig. 15. 400 MHz spectrum.

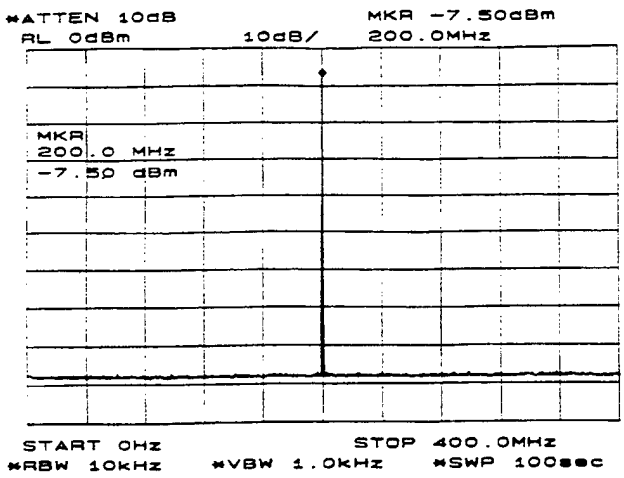


Fig. 16. 200 MHz spectrum.

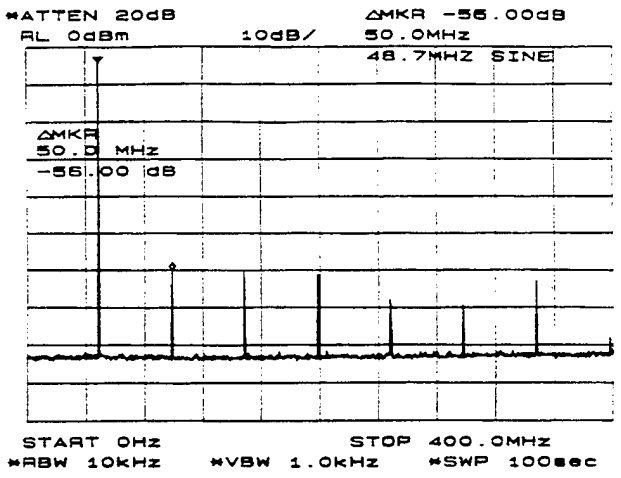


Fig. 17. 48.7 MHz spectrum.

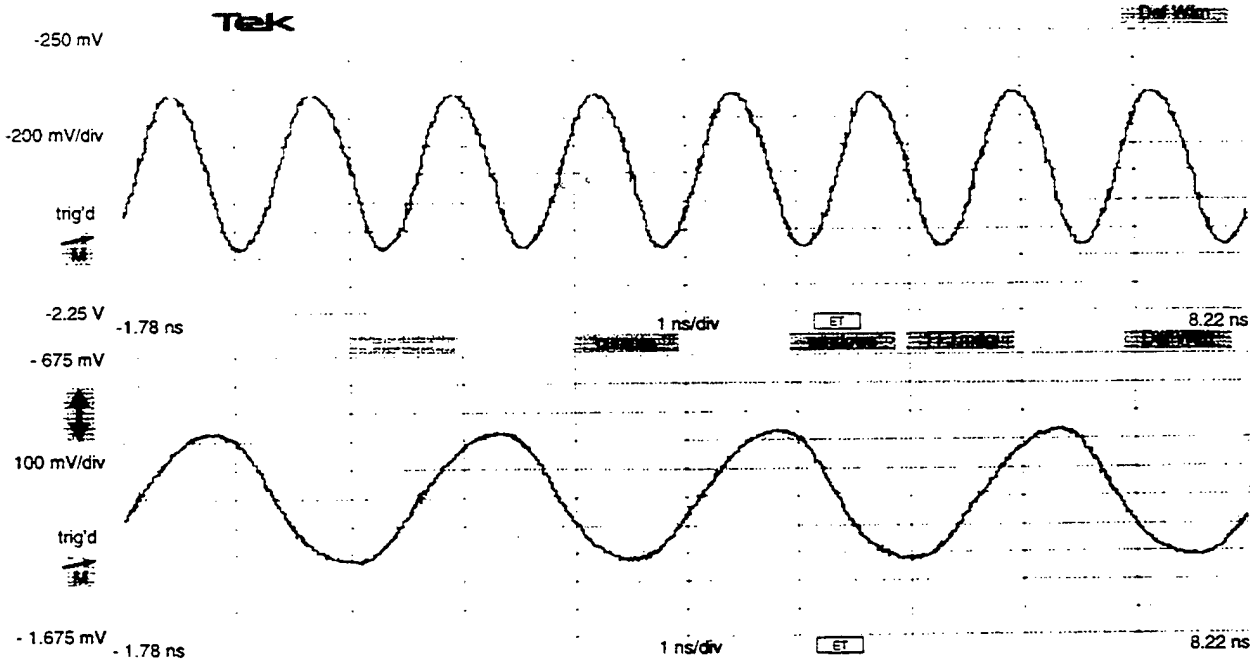


Fig. 18. 800 MHz clock and MSB of cosine from ECL output driver switching at 400 MHz.

IV. Fabrication and Test Results

Testing of the DDFS chip was accomplished by fabricating a testboard containing two GaAs 10-b D/A converters. The device used was the Rockwell RI61008, which has a maximum operating frequency of 1.2 GHz. The CMOS DDFS die is mounted in a 132 position, quad flat pack leaded ceramic package featuring controlled impedance 50 Ω lines for all signal paths. This package is installed in a low profile test socket utilizing wire button with plunger contact technology for a very low inductance and short electrical path interconnect to signal, power, and ground paths. The wire button contacts then mate with small gold plated pads located on microstrip lines that exist on the top layer of a six layer epoxy fiberglass (FR-4) circuit board that was designed and fabricated to evaluate the QDDFS4 chip. The other five layers of the circuit board were used to provide ac and dc ground planes, power and bias distribution paths, and control signal paths.

The block diagram of Fig. 14 illustrates the test setup. An HP8643A RF signal source was used to produce the reference clock. An HP8131A dual channel pulse generator reference was used in slave mode to amplify the reference clock to appropriate ECL levels. Analog outputs of the QDDFS4 and D/A converter combination were viewed using a TEK DSA602 digitizing signal analyzer. Spectrum measurements were per-

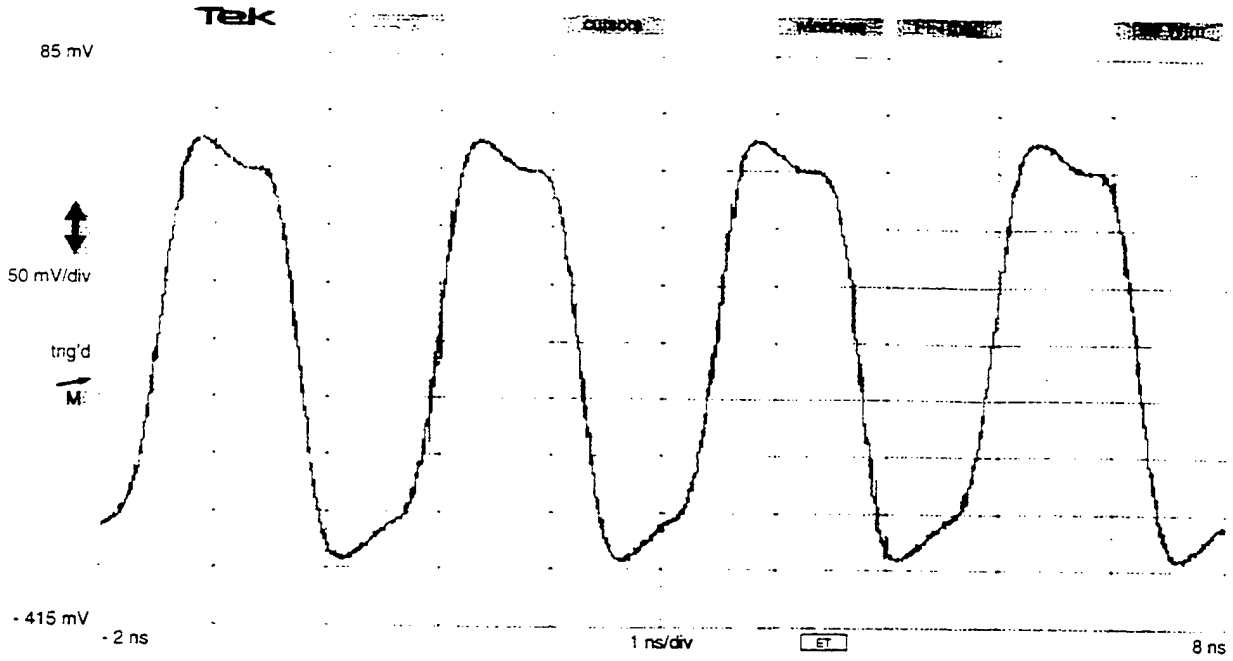


Fig. 19. 400 MHz output signal from D/A converter.

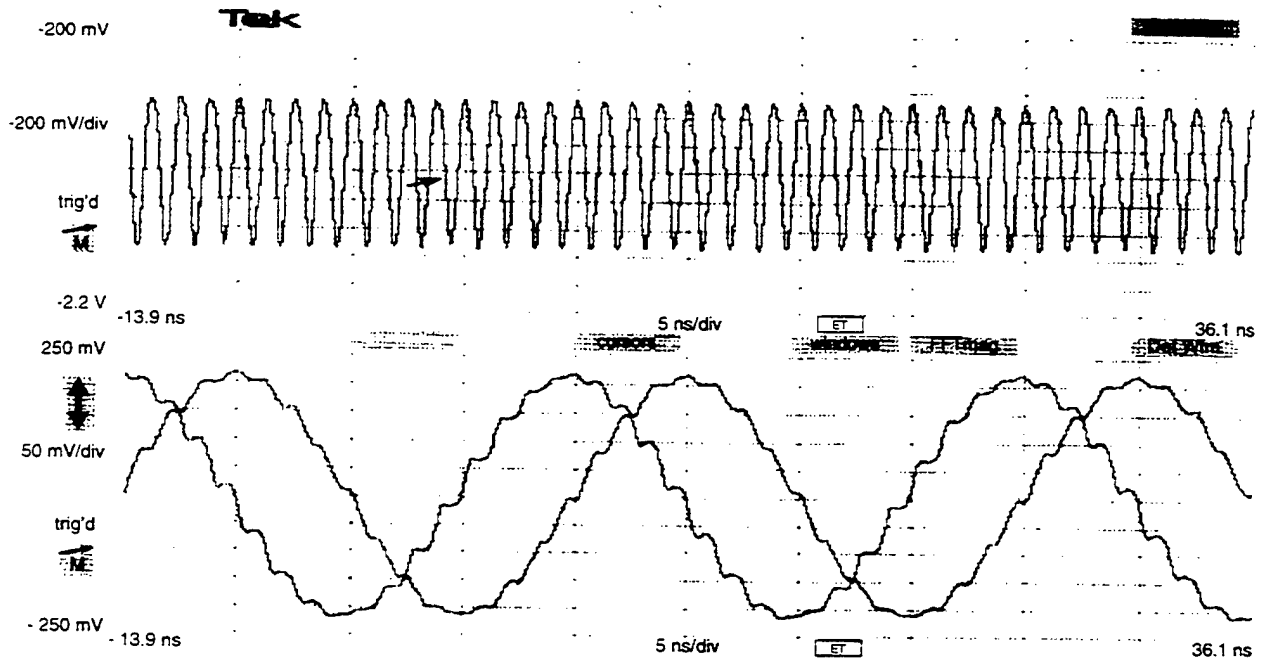


Fig. 20. Unfiltered 50 MHz output sine and cosine waveform from D/A converter.

formed with an HP8562A spectrum analyzer. All tests results presented here used a reference clock frequency of 800 MHz. Figs. 15, 16, and 17 illustrate spectrum plots of a 400 MHz, a 200 MHz, and a 48.7 MHz output sinewave, respectively. These plots span a wide frequency range to facilitate the viewing of harmonics produced by the nonlinearity of the D/A converter. The 400 MHz spectrum represents the case where $f_{out} = f_s/2$, which is the highest possible synthesizable frequency. The bottom trace of Fig. 18 illustrates the MSB of a cosine ECL output-driver switching at 400 MHz. The top trace is the corresponding reference clock at 800 MHz. The waveform is produced by setting the 32 b FCW of the QDDFS4 chip to equal 80 000 000(H), thus generating an $f_{out} = 400$ MHz. The loading of the output-driver is 50 Ω

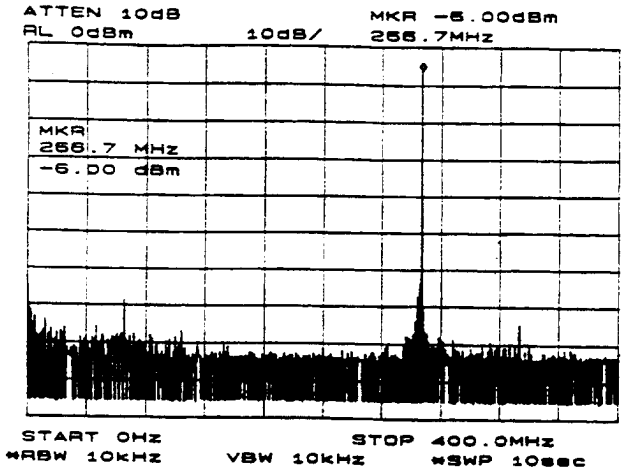


Fig. 21. Spectrum of $F_s/3 + \Delta$, where $\Delta = 65.1$ kHz, 400 MHz span.

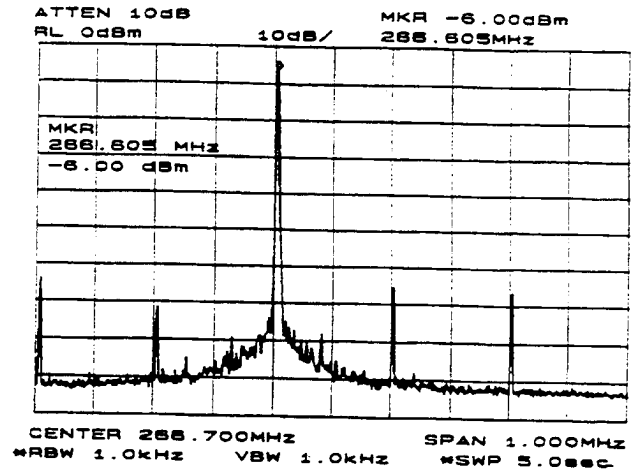


Fig. 22. Spectrum of $F_s/3 + \Delta$, where $\Delta = 65.1$ kHz, 1 MHz span.

in parallel with the loading of a 450 Ω TEK 6231 active probe used to sample the signal. Fig. 19 is the corresponding oscilloscope trace of the 400 MHz sinewave at the output of the D/A converter. The plug-in amplifier used has a 1 GHz bandwidth and thus partially filters the 400 MHz D/A converter output. Fig. 20 displays oscilloscope traces of a pair of unfiltered quadrature sine and cosine D/A converter output waveforms at 50 MHz. The signal trace above the quadrature signals is the 800 MHz reference clock. As shown in this figure, the sine and cosine waveforms are comprised of 16 discrete levels. Figs. 21 and 22 illustrate the QDDFS4 chip producing a $f_s/3 + \Delta$ fundamental where $\Delta = 65.1$ kHz. A frequency offset equal to Δ from $f_s/3$ produces a spreading of the aliased harmonic components spaced 3Δ apart. The worst case harmonic component is approximately 58 dB below the fundamental. The QDDFS4 produces a -84.3 dBc spectrally pure signal over its entire tuning range. Hence, the harmonic dibstortion in the analog outputs is primarily due to the D/A conversion process.

V. Conclusion

A quadrature digital synthesizer chip utilizing a parallel architecture has been designed that operates at 800 MHz and synthesizes -84.3 dBc spectrally pure sine and cosine digitized waveforms. This chip exhibits large bandwidth (dc to 400 MHz), high spectral purity, fast switching speed, and fine frequency resolution (0.188 Hz). The use of parallelism is an effective technique for increasing throughput in CMOS technology. Parallelism can be used successfully in CMOS processes simply because the yields are very high. Furthermore, effective use of parallelism provides a designer with the freedom to perform area versus speed trade-offs thus enabling production costs to be optimized for a specific targeted performance criterion. This technique is less effective in other high performance GaAs or silicon bipolar technologies since these fabrication processes are not able to match the yields of CMOS.

TABLE III
QDDFS4 Chip Specifications

Technology	0.8 μm HPCMOS26B, Triple Level Metal, N well Process
Maximum Clock Frequency	800 MHz
Switching Time	4 clock cycles or 5 ns (at 800 MHz f_{clk})
Tuning Latency	47 clock cycles
Tuning Bandwidth	400 MHz (at 800 MHz f_{clk})
Frequency Resolution	0.188 Hz (at 800 MHz f_{clk})
Worst Case Spurious	-84.3 dBc
Phase Modulation Wordlength	12 b
Output Word Length	12 b
Clock Input	Small Signal Swing Clock ± 0.5 V
Output Drivers	ECL-Compatible levels
Power Dissipation	3 Watts @5 V with I/Os switching (at 800 MHz f_{clk})
Power Efficiency	3.75 mW/MHz @5 V
Transistor Count	94,000 (including I/Os)
Die Size	7.2 mm x 7.9 mm
Core Size	5.9 mm x 6.7 mm

The QDDFS4 chip dissipates 3 W at 800 MHz operation thus providing a power efficiency of 3.75 mW/MHz. The power dissipation achieved is the lowest of any reported DDFS of the same speed performance class in any fabrication process. By taking advantage of sine and cosine symmetries, the size of the ROM lookup table is no larger than that of a DDFS, which only generates sine outputs. The QDDFS4 chip attains a fourfold speed increase over the previous fastest CMOS DDFS [6] and is the fastest 12 b output single-chip DDFS ever reported in any technology. Table III summarizes the chip specifications. A chip photomicrograph is shown in Fig. 23.

Acknowledgment

The authors wish to thank F. Lu for valuable discussions, E. Berg for work on the ROM, D. Kruse for register characterization, and R. Patel for laying out the printed circuit board. High speed testing was facilitated by equipment donations from Hewlett Packard and Tektronix. N. H. Sheng and D. Mehrotra of Rockwell International are gratefully acknowledged for technical assistance and for supplying samples of the RI61008 D/A converter.

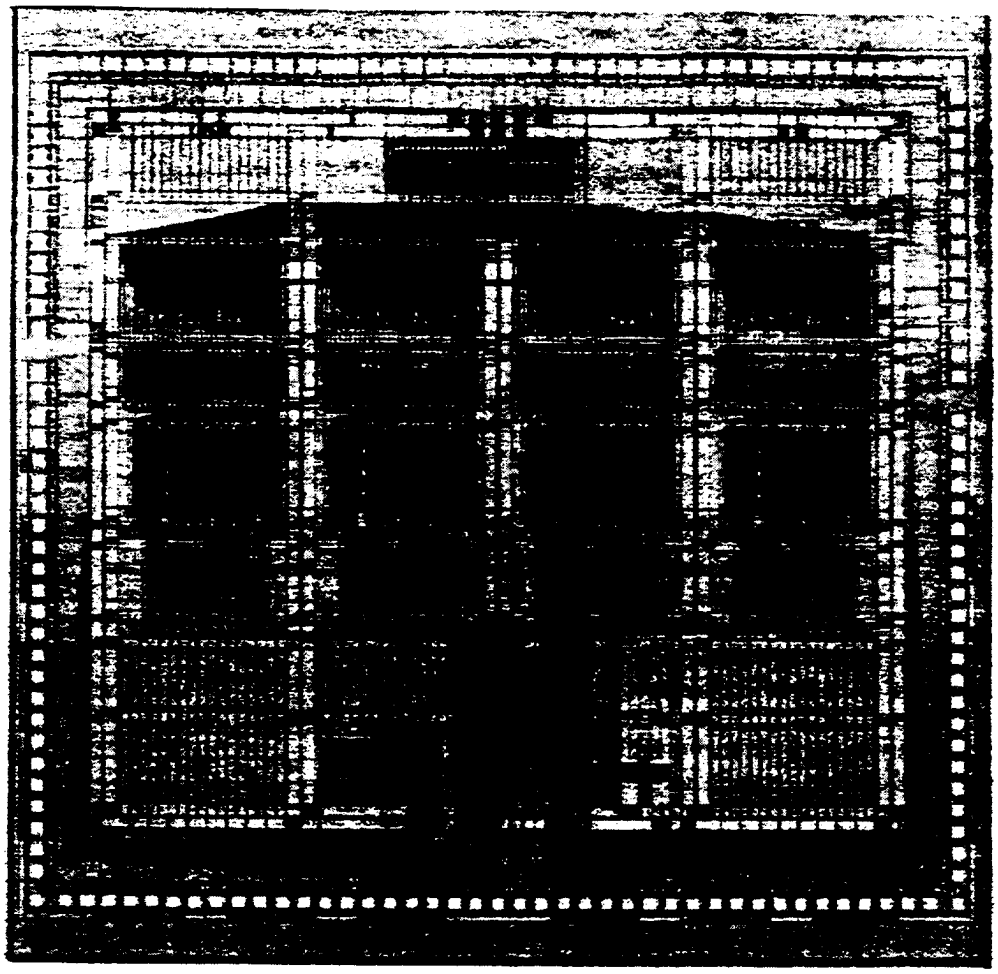


Fig. 23. QDDFS4 chip photomicrograph.

REFERENCES

[1] V. Manassewitch, *Frequency Synthesizers, Theory and Design*, 2nd ed. New York: Wiley, 1989.

[2] P. H. Saul and D. G. Taylor, "A high speed direct frequency synthesizer," *IEEE J. Solid-State Circuits*, vol. 25, no. 1, pp. 215–219, Feb. 1990.

[3] V. Andrews, C. T. M. Chang, J. D. Cayo, S. Sabin, W. A. White, and M. P. Harris, "A monolithic digital chirp synthesizer chip with I and Q channels," *IEEE J. Solid-State Circuits*, vol. 27, pp. 1321–1326, Oct. 1992.

[4] M. Thompson, "Low latency, high-speed numerically controlled oscillator using progression-of-states technique," *IEEE J. Solid-State Circuits*, vol. 27, pp. 113–117, Jan. 1992.

[5] R. Hassun and A. W. Kovalick, "Waveform synthesis using multiplexed parallel synthesizers," U.S. Patent no. 4, 454, 486, Hewlett Packard, June 12, 1984.

[6] L. K. Tan and H. Samueli, "A 200-MHz quadrature frequency synthesizer/mixer in 0.8 μm CMOS," *IEEE J. Solid-State Circuits*, vol. 30, pp. 193–200, Mar. 1995.

[7] H. T. Nicholas (III) and H. Samueli, "An analysis of the output spectrum of direct digital frequency synthesizers in the presence of phase-accumulator truncation," in *Proc. 41st Annual Frequency Control Symp. USERACOM*, Ft. Monmouth, NJ, May 1987, pp. 495–502.

[8] J. Yuan and C. Svensson, "High-speed CMOS circuit technique," *IEEE J. Solid-State Circuits*, vol. 24, no. 1, pp. 62–70, Feb. 1989.

[9] B. A. Chappell, T. I. Chappell, S. E. Schuster, H. M. Segmuller, J. W. Allan, R. L. Franch, and P. J. Restle, "Fast CMOS ECL receivers with 100-mV worst case sensitivity," *IEEE J. Solid-State Circuits*, vol. 23, no. 1, pp. 59–67, Feb. 1988.

[10] S. R. Meier, E. De Man, T. G. Noll, U. Loibl, and H. Klar, "A 2-μm CMOS digital adaptive equalizer chip for QAM digital radio modems," *IEEE J. Solid-State Circuits*, vol. 23, no. 5, pp. 1212–1217, Oct. 1988.

Loke Kun Tan was born in Kuala Lumpur, Malaysia, on April 2, 1964. He received the B.S. degree in electrical engineering from the University of Houston in 1987 and the M.S. and Ph.D. degrees also in electrical engineering from the University of California, Los Angeles in 1992 and 1995, respectively.

Since 1993, he has been at Broadcom Corporation where he works in the communications IC design group. He is currently involved in the design of QAM tranceivers for cable television applications. His research interests include digital signal processing, carrier synchronization, digital communications, and high performance CMOS circuit design.

Edward W. Roth was born in New York City, New York on December 23, 1954. He received the A.S.E.E. degree from West Los Angeles College, Culver City, CA, in 1981, the B.S. degree in business management from Pepperdine University, Los Angeles, CA in 1984, and the M.S. degree in procurement and acquisitions management from Northrop University, Los Angeles, CA in 1991.

Since 1992, he has been a member of the technical staff, and Laboratory Manager of the Integrated Circuits and Systems Laboratory at the University of California, Los Angeles. Prior to this, he was employed by the Hughes

Aircraft Company in Los Angeles, California from 1978 to 1985. During this time, he worked on the design and fabrication of microwave and millimeter wave components and systems. From 1985 to 1991, he was with the GigaBit Logic Co. (TriQuint Semiconductor) as a Project Engineer for High Frequency printed circuit board and test fixture design and development programs. From 1991 to 1992, he worked as a consultant designing microwave test fixtures for materials testing. His research interests include the areas of R.F., microwave, and millimeter wave communications and signal processing.

Gordon E. Yee was born in San Francisco, CA, in 1969. He received the B.S. degree in electrical engineering and computer sciences from the University of California, Berkeley, in 1991 and the M.S. degree in electrical engineering from the University of California, Los Angeles, in 1993.

Since 1993 he has been a Design Engineer at Cirrus Logic, Inc., Fremont, CA, where he is working in the area of graphics accelerators for personal computers. He is currently involved in the design of high-performance CMOS integrated circuits and VLSI architectures.

Henry Samueli (S'75–M'81) was born in Buffalo, NY, on September 20, 1954. He received the B.S., M.S., and Ph.D. degrees in electrical engineering from the University of California, Los Angeles in 1975, 1976, and 1980, respectively.

From 1980 to 1985, he was with TRW, Inc., Redondo Beach, CA where he was a Section Manager in the Digital Processing Laboratory of the Electronics and Technology Division. His group was involved in the hardware design and development of military satellite and digital radio communication systems. From 1980 to 1985, he was also a part-time instructor with the Electrical Engineering Department at UCLA. In 1985, he joined UCLA full time where he is currently a Professor with the Electrical Engineering Department. His research interests are in the areas of digital signal processing, communications systems engineering, and CMOS integrated circuit design for applications in high-speed data transmission systems. In 1988, he cofounded PairGain Technologies, Tustin, CA, a telecommunications equipment manufacturer, and in 1991, he cofounded Broadcom Corporation, Irvine, CA, a communications IC supplier.

Dr. Samueli is the recipient of the 1988/1989 TRW Excellence in Teaching Award of the UCLA School of Engineering and Applied Science and the Meritorious Paper Award of the 1991 Government Microcircuit Applications Conference.

A 1 GHz CMOS RF Front-End IC with Wide Dynamic Range

A. Rofougaran, J. Y-C. Chang, M. Rofougran,
S. Khorram, and A.A. Abidi

21st European Solid-State Circuits Conference Proceedings
pp. 250-253
September 1995.

This paper was presented at the *21st European Solid-State Circuits Conf.*, Lille, France, September 1995.
It appears on pages 250-253 of the conference proceedings.

A 1 GHz CMOS RF Front-End IC with Wide Dynamic Range

A. Rofougaran, J. Y-C. Chang, M. Rofougaran, S. Khorram, and A. A. Abidi

Integrated Circuits & Systems Laboratory
Electrical Engineering Department
University of California
Los Angeles, CA 90095-1594

Abstract

A 1 GHz Low Noise Amplifier and Mixer combination has been integrated in a standard 1μm CMOS process. The circuit is matched to 50Ω at the input, and drives a 50Ω load. Overall conversion gain is 22 dB, noise figure is 3.5 dB, and the IIP3 of the combination is +12 dBm. The fully balanced circuit drains 8 mA from 3V.

Introduction

There is now considerable interest in the possible use of CMOS ICs for RF applications at 900 MHz and above. Compared to silicon bipolar ICs, CMOS offers low fabrication cost, ready integration of analog and digital components, and what may not be readily apparent, some fundamental advantages in attainable dynamic range when a frequency mixer is required. Dynamic range is set at the lower end by device thermal noise, and limited at the upper end by device distortion. MOSFETs are superior in the latter respect, because their quadratic nonlinearity may be effectively suppressed by a balanced excitation to obtain very linear characteristics overall. Resistor degeneration may similarly linearize bipolar transistor circuits, but it does not offer the same flexibility as FETs do, and its effectiveness increases with quiescent power dissipation.

CMOS RF circuits are still at an early stage of exploration. In past work, we have shown how to build low-noise amplifiers in CMOS, which are tuned by large on-chip spiral inductors suspended above the substrate after a maskless post-fabrication etch [1]. Using a switched-capacitor subsampling technique for downconversion, we have demonstrated a 1 GHz downconversion mixer in 1-μm CMOS attaining a two-tone 3rd-order input-referred intercept point (IIP3) of +26 dBm, but whose noise figure is 18 dB due to aliasing of wideband noise [2]. Another 1.5 GHz continuous-time downconversion mixer uses the nonlinearity cancellation in cross-coupled FET resistors described above [3], to attain an IIP3 of +28 dBm with a 24 dB noise figure [4]. For these mixers to be usable in a sensitive receiver, they must be preceded by a high gain, low-noise RF amplifier (LNA). The wide mixer dynamic range is available when the LNA itself has a comparable dynamic range. Further, the LNA should be able to drive the amplified RF signal into the mixer. It is quite a challenge to meet these requirements with low current drain while operating at low voltages.

This paper describes a 1 GHz front-end, consisting of an LNA and mixer integrated on a 1-μm CMOS IC. The overall front-end, with a two-tone IIP3 of +12 dBm and noise figure of 3.5 dB, operates at 3V nominal and drains 8mA. While there is a continuous tradeoff between dynamic range and current drain, we note that the dynamic range at this current exceeds that of other similar state-of-the-art bipolar front-end ICs.

Circuit Description

Low Noise Amplifier

The LNA must provide a sufficiently high gain in the RF band of interest, it must have a low noise figure, and it must be well-matched at the input port to the source impedance, typically 50Ω for an antenna and passive preselect filter. When a noise figure less than 3 dB is sought, the RF input is applied to the gate of a suitably-sized FET, and input matching is obtained with inductively degeneration at the source and a series tuning inductor at the gate [5, 6]. If, on the other hand, a noise figure of 3dB or higher is acceptable, the RF input may be applied directly to the source of a common-gate FET

The UCLA Low-Power Wireless Transceiver Project is supported by the US Advanced Research Projects Agency, and by a consortium of semiconductor companies under the State of California MICRO Program. For more information, e-mail abidi@icsl.ucla.edu

whose $1/g_m$ is scaled to the matching resistance. This technique has previously been used in a 100 MHz bipolar RF IC [7]. Our LNA requires a balanced RF input, derived from a lossless, passive balun merged into the antenna. This is applied to the sources of two identical common-gate FETs (M1,M2), each loaded at the drain by an on-chip 50 nH spiral inductor (Fig.1). In this way, the overall noise figure is the same as if the single-ended antenna output is applied to one of the two FETs, yet this balanced LNA inherently rejects common-mode stray signals coupling into it at the input or through the substrate. One triode-biased PMOSFET attached to the other end of the inductors sets the nominal common-mode level of the LNA output equal to the gate bias voltage, VG.

FET C_{GS}, source junction capacitance, and the pad capacitance contribute a substantial parallel reactance to the FET input at 1 GHz. The sum of these capacitances is made to resonate at 1 GHz in parallel with a discrete off-chip 8 nH inductor, connected from the input pad to ground. This inductor also carries the FET bias current. All RF input pads are Metal 2 over grounded Metal 1, which shields these nodes from loss and excess noise that the spreading resistance of the underlying substrate would otherwise contribute [8]. With the FET size set at a given $(V_{GS}-V_t)$ by input matching, the total capacitance at its drain, including capacitive loading by the mixer, sets the tuning inductance. Let us say that the total drain capacitance $C \approx g_m/\omega_T$, where ω_T is the radian unity current-gain frequency at this bias. The LNA voltage gain is then set by the tuned circuit impedance at the resonant frequency, $\omega_0 = 2\pi$ Grad/s, as follows:

$$\text{Gain} = g_m Q^2 R_s = \left(\frac{g_m}{C}\right) \times \left(\frac{L}{R_s}\right) = 2\pi f_T \left(\frac{L}{R_s}\right)$$

where Rs is the series resistance of the inductor windings, and Q its quality factor. S-parameter based measurements of NMOSFET f_T at various gate biases (Fig.2) suggest that an acceptably high LNA voltage gain of 14 (23 dB) is obtained when the common-gate FETs are biased at $(V_{GS}-V_t)=0.6V$. For the 50 nH on-chip spiral inductor used, $Rs \approx 50\Omega$.

The relatively high bias point also improves the LNA linearity, because at the peaks of a large RF input the FETs may turn off or enter their triode region. However, it also means that the FETs operate at a low g_m/I, implying a higher current drain. The common-gate FETs with balanced input will inherently cancel the quadratic nonlinearity at their balanced output better than, say, a FET differential pair biased at a constant tail current.

Mixer

The mixer (Fig.3) consists of a pair of common-source FETs (M4, M5) acting as a *linear* (balanced) *V-to-I converter* for the amplified RF at the LNA output, cascaded by a *current-mode commutating switch* consisting of two differential pairs (M6-M9) cross-coupled at the outputs. At this point, the signal has been downconverted to the (low) intermediate frequency (IF). The commutated currents flow into a two-FET resistor load (M10, M11), which applies common-mode feedback from its center tap to bias the mixer with PMOS current sources. The IF voltage is buffered to an off-chip 50Ω load. The LNA bias voltages VC is adjusted to make the gate bias of M4, M5 equal to VG.

The commutating differential pairs require a large LO drive at the gate to fully switch the RF currents, and thereby they attain the maximum conversion gain. A sufficiently large LO signal also causes them to make a rapid excursion through their high gain region, thereby reducing the duty-cycle over which their own amplified noise contributes to the output [9]. A 1V ptp (5 dBm) sinewave is used in all the measurements. This gives the maximum RF-to-IF conversion gain of 2×.

The mixer output at IF is buffered by 960μm wide open-drain FETs into the off-chip 50Ω load presented by the measuring instruments. The buffer contributes a loss of roughly 2×, thus making the buffered mixer conversion gain close to 0 dB.

Experimental Results

The LNA-mixer combination is fabricated in a standard double-metal 1-μm CMOS process, using a p^+ substrate with a p-type epi layer. The LNA drains 5.2 mA, and the mixer 1.4 mA. The two output buffer FETs drain an additional 1.4 mA. The active area on the die is 1.3×2 mm, including the two on-chip spiral inductors (Fig.4). The substrate under the inductors is selectively removed on the fully fabricated die by a gas-phase isotropic etchant, which selectively attacks the bare silicon exposed through the eight circular windows surrounding each inductor. A rectangular pit roughly 150 μm deep results, which almost entirely removes the substrate capacitance, pushing the inductor self-resonance beyond 3 GHz. No extra mask steps need be added to the standard process. The inductor turns are built on Metal 2, with a Metal 1 return.

Power-splitters with anti-phase outputs create balanced signals from single-ended sources, which are then applied to the RF and LO ports. The balanced IF output is converted by a power-combiner into a single-ended signal for measurements. Measurements are made at a nominal 10 MHz IF by sweeping the RF and LO signal sources together with this offset frequency. This IF is well below the rolloff frequency of the mixer and its output buffer.

The LNA frequency response thus measured (Fig.5) shows good conformity to simulations. Tuning is almost exactly at 1 GHz, as designed, which validates the quality of the FET models, and of the simple three-element model we have previously used to describe the spiral inductor [1]. Unaccountable losses in the package, as well as instability at high gain settings caused by parasitics, make it difficult at this stage to reliably measure the peak gain. Nevertheless, it is within 2 dB of the expected value.

The input reflection parameter, s_{11}, determines the quality of the input match. Measurements show that the input port matches very well to 50Ω, attaining a minimum of -15 dB (Fig.6). However, owing to use of a smaller than optimum off-chip inductor at the input port, the minimum occurs at 1.5 GHz.

The nonlinearity of the *combined* LNA and mixer is measured with a two-tone test. The strength of the fundamental tones and 3rd-order intermodulation products generated by nonlinearities when the input is stimulated by two nearby frequencies at 1 GHz is measured at the IF output, and plotted against the total input signal strength (Fig.7). The two plots on a log scale have a relative slope of 1:3, and they extrapolate to an input-referred intercept point (IIP3) of +12 dBm. The 1-dB compression point of the overall gain is at an input of about -3 dBm.

The noise figure of the combined LNA and mixer attains a minimum of 3.5 dB at about 1 GHz (Fig.8). The simulated noise figure is 2.8 dB, due to the slightly lower noise in a FET inversion layer than in an equivalent $1/g_m$ ohmic resistor. The discrepancy is again likely due to small but unaccountable losses in the device package.

The DC offset at the mixer output, and particularly how it varies with parasitic coupling of the LO to the RF port, is very important when the mixer is used in a direct-conversion receiver [10]. The typical DC offset at the IF output port is about 1 mV, and in the extreme condition of turning the LO off from its nominal level, it is found to vary by only about 9μV.

Conclusions

This paper describes the first all-CMOS RF downconversion front-end operating at 1 GHz. It is worth comparing the measured results above with specifications of state-of-the-art silicon bipolar and GaAs RF ICs intended for the same application. Extracting data from a summary of these devices [11], we find that the LNA-mixer noise figure varies from 2.5 to 6 dB, gain varies from 15 to 24 dB, and the best IIP3 for a silicon device is -12 dBm, while the lowest is -22 dBm.The current drain of devices capable of operation at 3V varies from 5 to 15 mA.

In comparison, the front-end reported here, fabricated in a modest 1-μm CMOS digital IC technology offers a slightly higher noise figure than the ICs with the very best performance listed above, yet at comparable gain it has *far superior linearity*. It drains 8 mA from 3V, yet offers a fully balanced input and output, which no commercial product does. This exceptional performance results from exploiting the innate characteristics of MOSFETs for the task, and from a low-cost maskless technology to fabricate large-value on-chip spiral inductors. With these design methods, CMOS may very well become the technology of choice for RF receivers.

References

1. J.Y.-C. Chang, A.A. Abidi, and M. Gaitan, "Large Suspended Inductors on Silicon and their use in a 2-μm CMOS RF Amplifier", *IEEE Electron Device Letters*, v. 14, no.5, pp. 246-248, 1993.
2. P.Y. Chan, A. Rofougaran, K.A. Ahmed, and A.A. Abidi, "A Highly Linear 1-GHz CMOS Downconversion Mixer", in *European Solid-State Circuits Conf.*, pp. 210-213, Sevilla, Spain, 1993.
3. J. Crols and M. Steyaert, "A Full CMOS 1.5GHz Highly Linear Broadband Downconversion Mixer", in *European Solid State Circuits Conf.*, pp. 248-251, Ulm, Germany, 1994.
4. J. Crols and M. Steyaert, "A Fully Integrated 900 MHz CMOS Double Quadrature Downconverter", in *Int'l Solid State Circuits Conf.*, pp. 136-137, San Francisco, 1995.
5. R.E. Lehmann and D.D. Heston, "X-Band Monolithic Series Feedback LNA", *IEEE Trans. on Microwave Theory & Techniques*, v. MTT-33, no.12, pp. 1560-1566, 1985.
6. E. Heaney, F. McGrath, P. O'Sullivan, and C. Kermarrec, "Ultra Low Power Low Noise Amplifiers for Wireless Communications", in *GaAs IC Symp.*, pp. 49-51, San Jose, CA, 1993.
7. T. Okanobu, H. Tomiyama, and H. Arimoto, "Advanced Low-Voltage Single Chip Radio IC", *IEEE Trans. on Consumer Electronics*, v. 38, no.3, pp. 465-475, 1992.
8. N. Camilleri, J. Kirschgessner, J. Costa, D. Ngo, and D. Lovelace, "Bonding Pad Models for Silicon VLSI Technologies and their effects on the Noise Figure of RF NPNs", in *Microwave & Millimeter-Wave Monolithic Circuits Symp.*, pp. 225-228, San Diego, CA, 1994.
9. R.G. Meyer and W.D. Mack, "A 1-GHz BiCMOS RF Front-End IC", *IEEE J. of Solid-State Circuits*, v. 29, no.3, pp. 350-355, 1994.
10. A.A. Abidi, "Direct-Conversion Radio Transceivers for Digital Communications", in *Int'l Solid State Circuits Conf.*, pp. 186-187, 363-364, San Francisco, 1995.
11. TriQuint Semiconductor, TQ9203J Downconverter Application Note. 1994, p. 7.

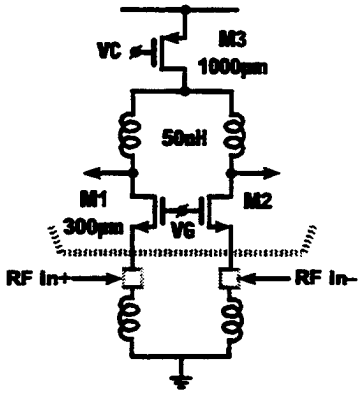


Fig.1 Circuit Diagram of Low-Noise Amplifier. Balanced RF input is derived from antenna, and output connects directly to mixer.

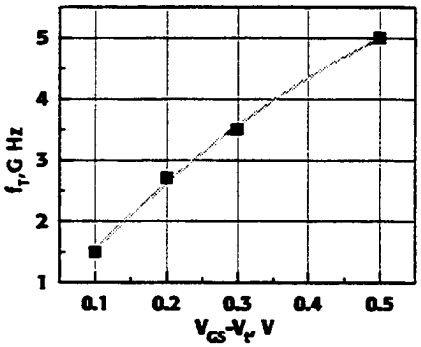


Fig.2 Measured dependence of f_t of 1-μm NMOSFETs on gate drive.

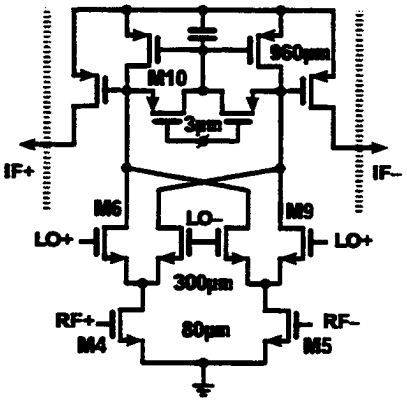


Fig.3 Circuit Diagram of Mixer. Gain is controlled by bias on PMOS load (M10, M11). Output is taken open-drain.

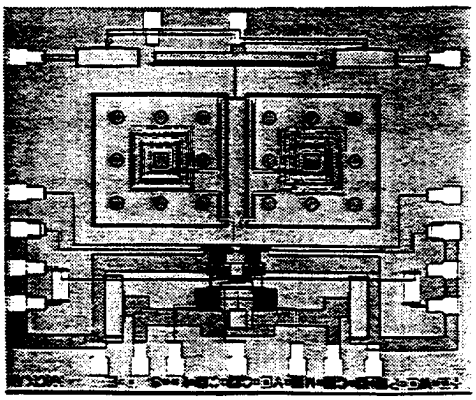


Fig.4 Die photo of LNA and mixer. Substrate under 50 nH spiral inductors is removed after fabrication to eliminate capacitance.

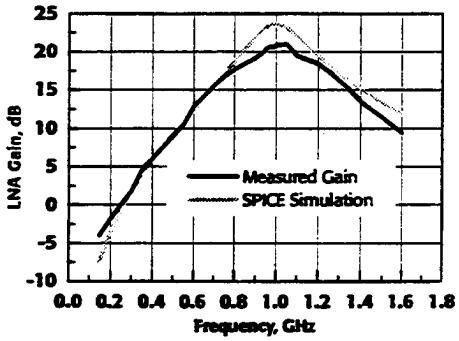


Fig.5 Measured LNA gain and frequency response compares well with SPICE simulations.

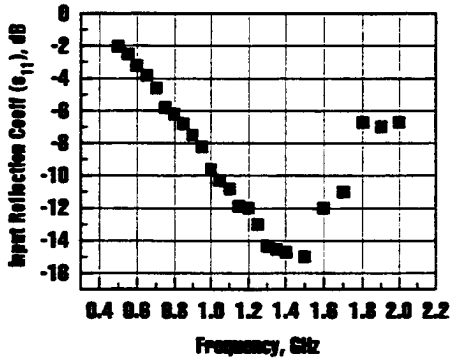


Fig.6 Measured matching at RF input port, as determined by s_{11}. Non-optimal off-chip inductor causes slight mistuning

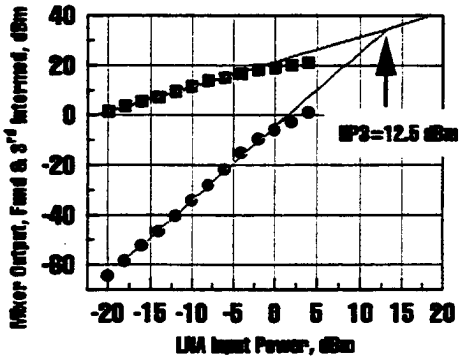


Fig.7 Measured nonlinearity of LNA and mixer with a two-tone RF input.

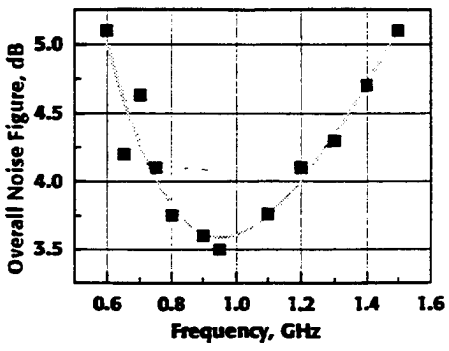


Fig.8 Measured Noise Figure of LNA/Mixer combination.

Analysis of a Square Helix Applicable to the Personal Satellite Communication Handset

J.S. Colburn and Y. Rahmat-Samii

Proceedings of the IEEE Antennas and Propagation society International Symposium
vol. 3, pp.1418-1421
June 1995

Analysis of a Square Helix Applicable to the Personal Satellite Communication Handset

Joseph S. Colburn* and Yahya Rahmat-Samii
Department of Electrical Engineering
University of California, Los Angeles
Los Angeles, CA 90024-1594

I. INTRODUCTION In the past few years there has been considerable work performed regarding antennas applicable to landbased wireless personal communication systems. This work has consisted of the design of small unobtrusive antennas [1] and studies on the effects of both the human operator on the transceiver antenna's performance and the transceiver's radiation on the human operator for the 915 MHz mobile communication channel [2]. This research concluded that the presence of the human operator has a significant impact on the antenna's performance – including loss in efficiency, perturbation of the input impedance and distortions of the far-field radiation pattern – and should be included in any realistic analysis.

Recently there has been significant interest in developing personal satellite communications as a means of providing very large areas of wireless communications coverage [3]. From the antenna design point of view these links offer new challenges in the way of requirements for circular polarization, slightly more directive radiating elements and higher operating frequencies (L and S band). Again the presence of the human operator must be contended with in this scenario.

In this paper, initial results from the study of one candidate antenna for a personal satellite transceiver handset, a square thin-wire helix, are presented with the intent to study the possible effects the presence of the operator will have on the radiation characteristics of circularly polarized antennas. The Method of Moments (MOM) is used to perform a full-wave analysis of the square thin-wire helix mounted on a small conducting box, and its performance is shown to be similar to a standard (circular) helical antenna. In addition, the geometry of the square helical antenna readily allows it to be simulated with a Cartesian grid Finite-Difference Time-Domain (FDTD) code that can also incorporate a model of a human head in the analysis. The effect the head model has on the computed far-field pattern and axial ratio will be presented.

II. ANALYSIS Figure 1 (a) is a schematic of a structure analyzed with MOM, a standard 4-turn helix (diameter = 0.306 λ, pitch = 0.24 λ) mounted on a 0.48λ x 0.48λ x 0.80λ conducting box. Figure 1 (b) illustrates a modification to the standard helix structure shown in Figure 1 (a), where the thin-wire element is constructed of only straight wire segments. All the vertically orientated wire segments are 0.06λ in length and all the horizontally orientated wire segments are 0.24λ in length. This square helix structure has the potential advantages of easier construction, a geometry easier to integrate with the

0-7803-2719-5/95/$4.00 © 1995 IEEE

handset and can be simulated with a Cartesian grid based FDTD code.

Figures 2 (a) and (b) are plots of the MOM computed far-field pattern in terms of right-hand circularly polarized (RHCP) and left-hand circularly polarized (LHCP) components, and axial ratio (in dB, 20log(major axis/minor axis) respectively for the standard helical antenna structure illustrated in Figure 1 (a). Figures 3 (a) and (b) are similar plots for the square helix structure shown in Figure 1 (b). From these plots it can be seen that both these structures have fairly similar characteristics, although the square helix structure has slightly more directivity, a significantly lower back lobe and better axial ratio performance.

Figures 4 (a) and (b) are plots of the computed far-field pattern and axial ratio for the same square helix structure using a Cartesian grid based FDTD code. A comparison of Figures 3 and 4 shows good correlation of the square helix structure computed performance between the MOM and FDTD codes.

In [2] a computational head model was included in the FDTD simulations in order to ascertain the impact the user's presence has on the radiation characteristics of the transceiver. In this work, a similar computational phantom head was used to help in the characterization of the personal satellite communication transceiver in close proximity to a user. Although the same structure and approximate head size were used, the electrical parameters of the head model were changed to those of head tissue at 1.8 GHz [4], see Figure 5. Figure 6 illustrates the computational head model positioned near the transceiver handset. At 1.8 GHz, the FDTD grid size of 0.02λ is 3.33 mm and the size of the transceiver box is 8 cm x 8 cm x 13.33 cm. In the computations performed, the handset and head model were spaced one cell apart.

Figures 7 (a) and (b) are plots of the FDTD computed far-field pattern and axial ratio for the square helix structure positioned next to the computation phantom head seen in Figure 6. The computations indicate approximately 13 percent of the total delivered power is lost in the head model, which is less than the losses reported in [2] and is mainly attributed to the greater distance between the head model and the antenna element. Besides the loss in radiated power due to absorption in the head model, it can be seen in Figures 7 (a) and (b) that there are significant distortions of both the computed far-field pattern and axial ratio due to the prescene of the head model.

Acknowledgement This work was supported in part by ARPA Contract #DAAB07-93-C-C501.

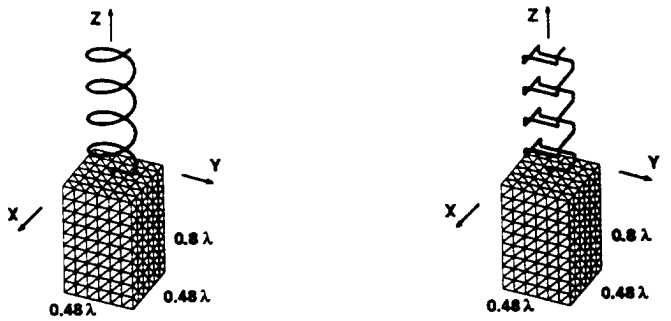
REFERENCES

[1] M. A. Jensen and Y. Rahmat-Samii, *IEEE Trans. Antennas Propagat.*, 42:1106–1113, 1994.

[2] M. A. Jensen and Y. Rahmat-Samii, *Proc. of the IEEE,* 83:7–17, 1995.

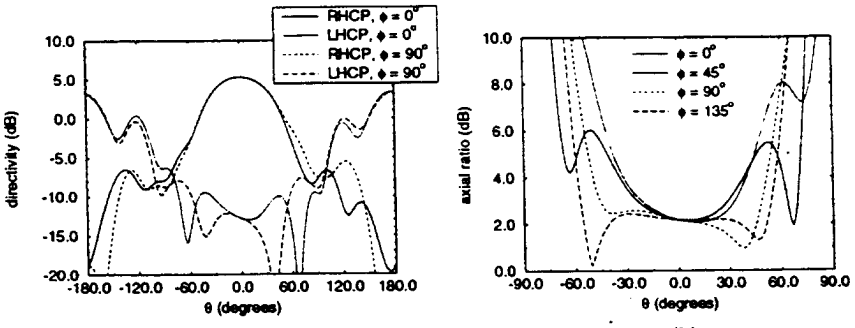
[3] W. W. Wu, E. F. Miller, W. L. Pritchard and R. L. Pickholtz, *Proc. of the IEEE,* 82:1431–1448, 1994.

[4] P. J. Dimbylow and O. P. Gandhi, *Phys. Med. Bio.*, 36:1075-1089, 1991.



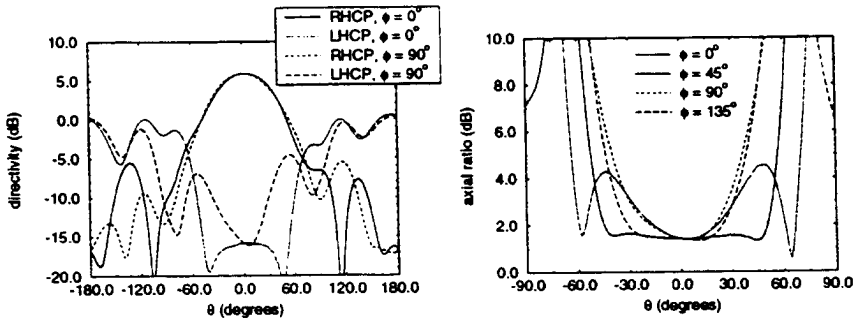
(a) (b)

Figure 1: Schematics of the structures analyzed: (a) standard helix (diameter = 0.306λ, pitch = 0.24λ) and (b) square helix; both on a 0.48λ x 0.48λ x 0.8λ conducting box.



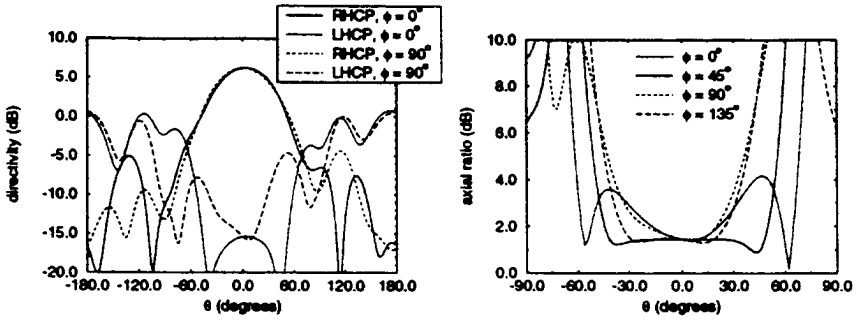
(a) (b)

Figure 2: MOM computed: (a) far-field pattern and (b) axial ratio; for the helix structure illustrated in Figure 1 (a).



(a) (b)

Figure 3: MOM computed: (a) far-field pattern and (b) axial ratio; for the square helix structure illustrated in Figure 1 (b).



(a) (b)

Figure 4: FDTD computed: (a) far-field pattern and (b) axial ratio; for the square helix structure illustrated in Figure 1 (b).

tissue	Permittivity	Conductivity (S/m)
Bone	8.0	0.15
Skin	32.0	0.57
Muscle	56.0	1.76
Brain	53.0	1.58
Humour	74.0	2.27
Lens	42.0	1.19
Cornea	51.0	2.29

Figure 5: Electrical characteristics of material in the computational head model at 1.8 GHz [4].

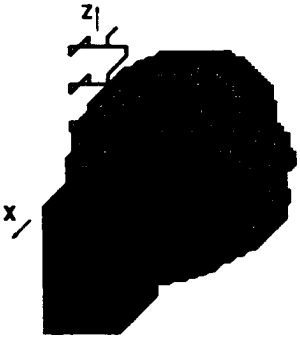
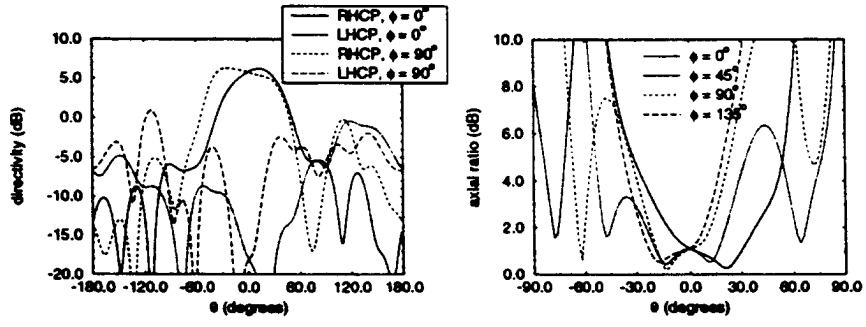


Figure 6: Square helical antenna on handset positioned near the computational head model.



(a) (b)

Figure 7: FDTD computed: (a) far-field pattern and (b) axial ratio; for the square helix structure near the head model illustrated in Figure 6.

A CMOS Limiting Amplifier and Signal-Strength Indicator

S. Khorram, A. Rofougaran, and A.A. Abidi

1995 Symposium on VLSI Circuits Digest of Technical Papers
pp. 95-96
June 1995

A CMOS Limiting Amplifier and Signal-Strength Indicator

S. Khorram, A. Rofougaran, and A. A. Abidi

Integrated Circuits & Systems Laboratory
Electrical Engineering Department
University of California
Los Angeles, CA 90024-1594

Introduction

Mobile radio handsets require an on-chip circuit to measure the received signal-strength. This information is sent to the base station to regulate the transmitted power level, and to determine cell-handoff. As strong attenuation and fading may cause the received signal power to span 6 to 8 decades, the measurement circuit must have a semi-logarithmic input-output characteristic to encompass this dynamic range.

The design of logarithmic amplifiers is now a well-established, although specialized, art. Unlike older implementations, these amplifiers no longer derive a logarithmic characteristic from the PN junction I-V characteristic; instead, they use the *successive-detection* architecture [1], wherein the sum of equally-weighted taps along a cascade of identical clipping amplifiers approximates the logarithm as a piecewise-linear function (Fig.1). A DC measure of the amplitude of a bipolar signal such as a sinewave is obtained by rectifying each tap, and lowpass filtering the sum of the rectified taps. These measurements will be accurate to the degree the actual amplifier input-output characteristic conforms to the ideal logarithm, and how stable it remains over the operating temperature range. Both these qualities are captured by the deviation of the actual characteristics from the ideal logarithm, specified as a gain ripple in dB.

Although all commercially available monolithic log amps today are bipolar ICs [2-5], CMOS is equally well-suited to implement the successive-detection architecture. We report here on the design and performance of such a logarithmic amplifier, which is part of a monolithic all-CMOS spread-spectrum 900 MHz wireless transceiver [6]. In the intended use, a received 160 kb/s binary-FSK signal is amplified at RF, directly downconverted to DC, and applied to the logarithmic amplifier after channel-select filtering. The amplifier provides two useful outputs. First, the limited output from the cascade of clipping amplifiers contains the data encoded as signal *phase* in the zero-crossings. This is directly applied to a correlating 1-bit digital detector for binary-FSK [6], without need for AGC. Second, the circuit produces a logarithmic signal-strength measurement to an accuracy of 1 dB over a 80 dB dynamic range, which a slow 7-bit A/D converter digitizes for uplink to the base-station.

The salient features of this logarithmic amplifier are: operation to 2.7-V supply, 0.75 mA current drain with a 5 MHz bandwidth, 80 dB dynamic range, 84 dB limiting gain, and implementation in 1-μm CMOS.

Research supported by ARPA, Rockwell International, Texas Instruments, Harris Semiconductor, and the State of California MICRO Program.

Circuit Design

The useful dynamic range of a successive-detection amplifier is limited on the upper end when the input causes the first stage in the cascade to clip, and on the lower end when all stages are in the linear region. In practice, if the amplifier noise causes the last stage to limit, it will determine the lower end. The amplifier characteristics will accurately conform to the logarithm over this dynamic range when a large number of clipping amplifier blocks, each with a relatively small gain, is cascaded. For instance, with a gain of 10 dB per stage and an ideal rectifier, the ripple in the DC logarithmic characteristic is 0.3 dB [4]. We use a cascade of 7 direct-coupled, balanced clipping amplifiers, each with 12 dB gain, and sum the output of 8 rectifiers along the cascade (Fig.1). Each stage consists of an NMOS differential pair with NMOS loads, and its gain, therefore, depends mainly on device-ratio, and to a large extent is process- and temperature-independent (Fig.2). The input and output common-mode levels are the same. A four-quadrant MOS multiplier rectifies the balanced signal at each tap in the cascade (Fig.3). Single-ended output currents taken from all 8 rectifiers are summed into an on-chip 10kΩ polysilicon resistor, and filtered by an off-chip 1-μF capacitor. All circuits are designed for operation at 2.7 V supply.

A well-designed bias source is important in attaining stable logarithmic characteristics. An on-chip replica circuit generates bias voltages for the current sources in the amplifier and rectifier (Fig.3).

Experimental Results

An off-chip *RC* circuit with a cutoff frequency of 3 Hz is connected in negative feedback between the limiting output and amplifier input, to stabilize the bias and suppress DC offsets. All measurements are made with a variable-amplitude 455 kHz sinewave capacitively-coupled into the amplifier. This is the standard 2nd IF in FM cellular telephones, and a sufficiently high frequency to cover the baseband spectrum of 160 kb/s binary-FSK.

The amplifier output is observed to limit on its own noise, as is customary in limiting amplifiers of high sensitivity. The signal-strength output conforms well to the logarithm over more than 80 dB, with a minimum detectable signal of –80 dBm (32 μV amplitude), which is also close to the equivalent input noise voltage (Fig.4). A ripple of ±0.5 dB is observed at all tempera-

tures, but owing to a small temperature-dependence of gain, the error curve acquires a slope which exacerbates the total error to ±1 dB at 85°C relative to the room-temperature characteristic (Fig.4(b)). This is acceptable for the intended application, and a considerable improvement over the ±3 dB ripple in a previously published CMOS logarithmic amplifier [7].

A ±1V ptp differential square-wave with well-defined zero-crossings is observed at the limiting output across the entire dynamic range. The group-delay through the cascade must also stay relatively constant over this range [5], otherwise the detected SNR, which is sensitive to phase of the received signal, will suffer. We measure this to vary by less than 100 ns across all amplitudes (Fig.5), which the resultant 16° peak phase-error will cause only a 0.2 dB loss in SNR.

The 1-μm CMOS N-well IC occupies an active area of 1 sq mm, and drains a total current of 0.75 mA from 2.7 V.

References

1. R.S. Hughes, *Logarithmic Amplification with Application to Radar and EW.* Dedham, MA: Artech House, 1986.
2. D. Anderson and R.J. Zavrel, RF ICs for Portable Communications Equipment, in *Electronic Components and Applications.* v. 7, no.1, January 1985, pp. 37-44.
3. Motorola, MC3362 Low-Power Dual Conversion FM Receiver, in *Communications Device Data Book.* 1992, pp. 2/60-2/63.
4. Analog Devices, AD640 DC-Coupled Demodulating 120 MHz Logarithmic Amplifier, in *Special Linear Reference Manual.* 1992, pp. 3/31-3/46.
5. Analog Devices, IF ICs Offer 3V Operation for Digital and Analog Cellular Standards, in *Microwave J.* v. 37, no.10, October 1994, pp. 144-148.
6. J. Min, A. Rofougaran, H. Samueli, and A.A. Abidi, "An All-CMOS Architecture for a Low-Power Frequency-Hopped 900 MHz Spread-Spectrum Transceiver", in *Custom IC Conf.*, pp. 379-382, San Diego, CA, 1994.
7. K. Kimura, "A CMOS Logarithmic Amplifier with Unbalanced Source-Coupled Pairs", *IEEE J. of Solid State Circuits,* v. 28, no.1, pp. 78-83, 1993.

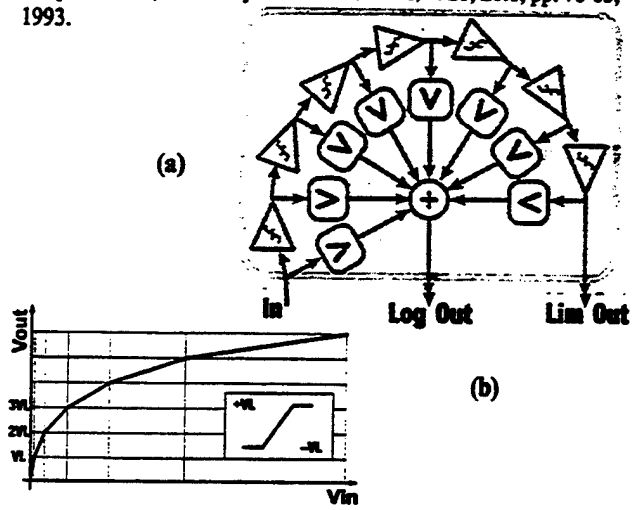


Fig.1. (a) Principle of successive-detection logarithmic amplifier. Yields both logarithmic and limited outputs. (b) Piecewise-linear approximation to log function. VL is clipping-level of each amplifier.

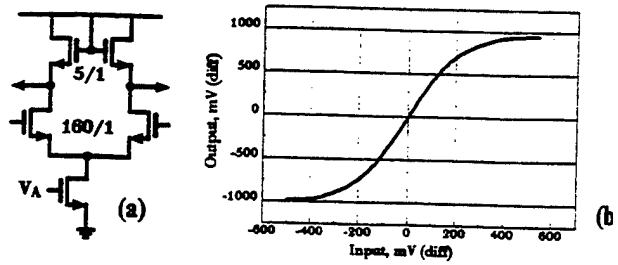


Fig.2. (a) Single clipping amplifier, and (b) its measured input-output characteristic.

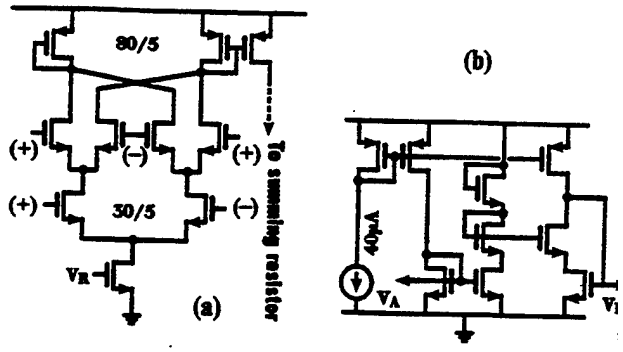


Fig.3. (a) Single full-wave rectifier. (b) Replica-bias circuit.

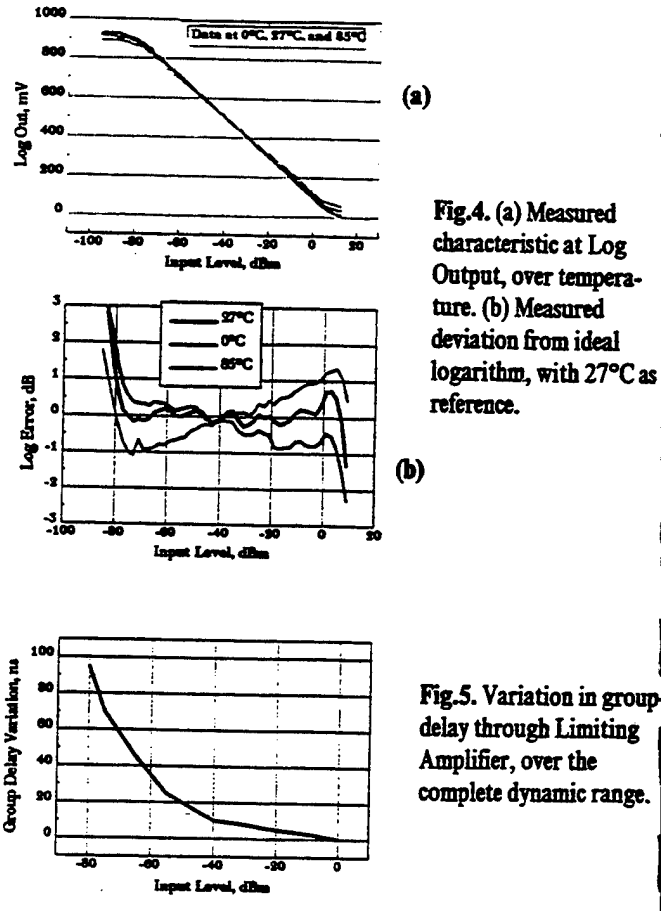


Fig.4. (a) Measured characteristic at Log Output, over temperature. (b) Measured deviation from ideal logarithm, with 27°C as reference.

Fig.5. Variation in group delay through Limiting Amplifier, over the complete dynamic range.

A Low-Power Correlation Detector for Binary FSK Direct-Conversion Receivers

J. Min, H.-C. Liu, A. Rofougaran, S. Khorram,
H. Samueli, and A.A. Abidi

Electronics Letters
vol. 31, no. 13, pp. 1030-1032
June 1995

16bit fixed point DSP and carried out experiments in fast Rayleigh fading channels without additive white Gaussian noise. The modulation method is π/4-shift DQPSK and the data rate is 42kbit/s. The GCA and the RSSI both have a 50dB dynamic range. The symbol synchronisation is hard-wired. Fig. 2 shows the input-output performance of the RSSI with the A/D converter and the input-output performance of the GCA with the D/A converter, respectively. From this Figure, we can see that there is a mismatch between the RSSI performance and the GCA performance. This mismatch can cause the performance degradation of the FF-AGC.

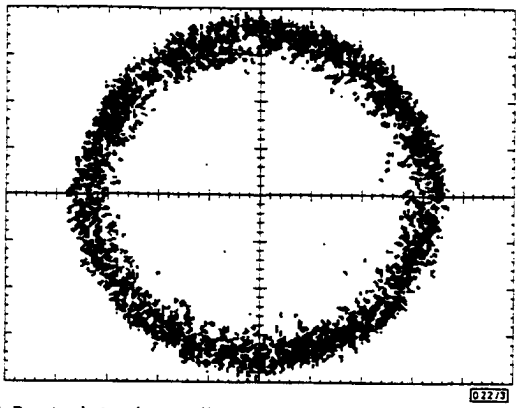


Fig. 3 *Received signal constellation suppressed by FF-AGC, f_0 = 80Hz*

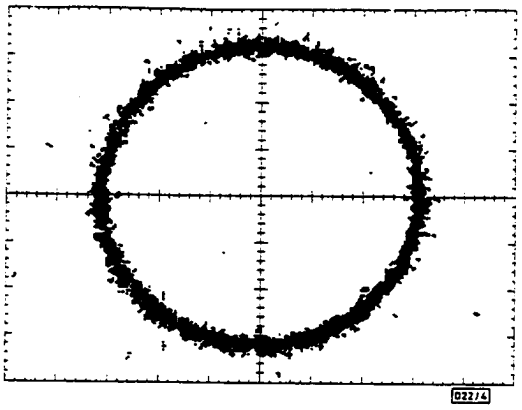


Fig. 4 *Received signal constellation suppressed by FF/FB-AGC, f_0 = 80Hz*

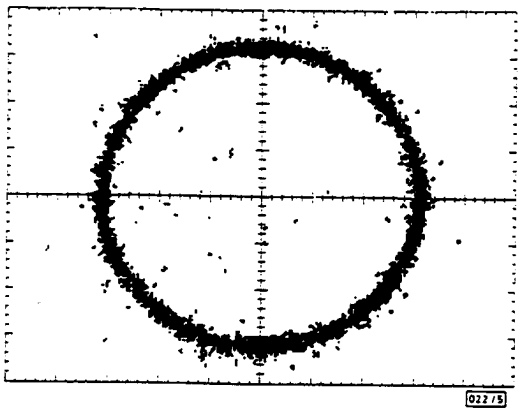


Fig. 5 *Received signal constellation suppressed by FF-AGC, f_0 = 160Hz*

Fig. 3 shows the received signal constellation sampled at symbol timing, i.e. $x(k)$ suppressed by the FF-AGC which corresponds to $\lambda = 0.75$ and $\mu = 0$ in a Rayleigh fading with a maximum Doppler frequency (f_D) of 80Hz. We can see that the received signal constellation suppressed by the FF-AGC has ~2.4dB deviation owing to the mismatch between the RSSI performance and the GCA performance or the lag error of the smoother response for fast variations. Fig. 4 shows the received signal constellation suppressed by the FF/FB-AGC with $\lambda = 0.75$ and $\mu = 0.1$ in a 80Hz Rayleigh fading channel. We can see that the proposed structure has a smaller deviation (1.4dB) compared with the result obtained using the FF-AGC. This is because the feedback control part of the FF FB-AGC mitigates the degradation owing to the mismatch or the lag error. Fig. 5 shows the received signal constellation suppressed by the FF/FB-AGC with $\lambda = 0.5$ and $\mu = 0.2$ in a 160Hz Rayleigh fading channel. We can see that the proposed structure maintains a small deviation (1.4dB) by setting smaller λ and larger μ values. From the experimental results, we see that the FF/FB-AGC improves the suppression performance compared to the FF-AGC, in fast Rayleigh fading channels.

Conclusion: We have proposed an FF/FB-AGC and examined the suppression performance. We have shown that an FF/FB-AGC can provide better suppression than a conventional FF-AGC in fast Rayleigh fading channels.

22 May 1995
Electronics Letters Online No: 19950736

M. Fujii, M. Nakamura and T. Ohsawa (*Fujitsu Laboratories Ltd., 1015 Kamikodanaka, Nakahara-ku, Kawasaki 211, Japan*)

N. Kawaguchi (*Fujitsu Tohoku Digital Technology Ltd., 2-3-9 Shinyokohama, Kouhoku-ku, Yokohama 222, Japan*)

References

1 MCGEEHAN, J.P., and BURROWS, D.F.: 'Performance limits of feedforward automatic gain control in mobile radio receivers', *IEE Proc. F*, 1981, **128**, pp. 385–392
2 SLADEN, J.P.H., and MCGEEHAN, J.P.: 'The performance of combined feedforward AGC and phase-locked AFC in a single sideband mobile radio receiver', *IEE Proc. F*, 1984, **131**, pp. 437–441
3 LINDE, L.P.: 'AGC strategy for adaptive digital modems employing cyclic channel estimation in an HF frequency hopping scenario', *IEE Proc. I*, 1991, **138**, pp. 185–188
4 MEYR, H., and ASCHEID, G.: 'Synchronisation in digital communications, Vol. 1' (John Wiley & Sons, 1989)

Low power correlation detector for binary FSK direct-conversion receivers

J. Min, H.-C. Liu, A. Rofougaran, S. Khorram, H. Samueli and A.A. Abidi

Indexing terms: Radio receivers, Frequency shift keying

A multiplierless binary FSK detector with 82dB of dynamic range is presented. The proposed detector is well suited to low power direct-conversion receivers used in wireless communications systems employing FSK modulation.

Introduction: Frequency-shift keying (FSK), often in continuous-phase format, is widely used in wireless communication systems such as pagers and frequency-hopped transceivers [1]. The modulated waveform has a constant envelope and narrow power spectra. At the receiver, noncoherent detection is often employed to reduce the hardware complexity. Direct-conversion receivers also have been gaining much attention recently for portable communications applications where low power is a key requirement. A direct-conversion architecture receiver translates the received signal from RF directly to I-Q baseband signals for quadrature detection. This eliminates image-reject filters and other IF components, and thus enables a complete monolithic transceiver in a one-chip or two-chip solution. Therefore, it is critical to find an efficient FSK baseband detector for such a direct-conversion receiver. Conventional IF FM detectors such as limiter-discriminator detectors are not suitable. Some new baseband binary FSK detectors for DECT and radio paging systems have been proposed [2 – 4]. It is well known that the optimum FSK detector is a correlation detector [5]. However, this detector is not often used in practice owing to the complexity of the required circuits. We propose a simple multiplierless binary FSK correlation detector for use in a direct-conversion receiver. The quadrature input signals are first hard-limited using a limiting amplifier with a high

dynamic range, thus eliminating the need for a multi-bit analogue-to-digital converter (ADC) or automatic gain control (AGC).

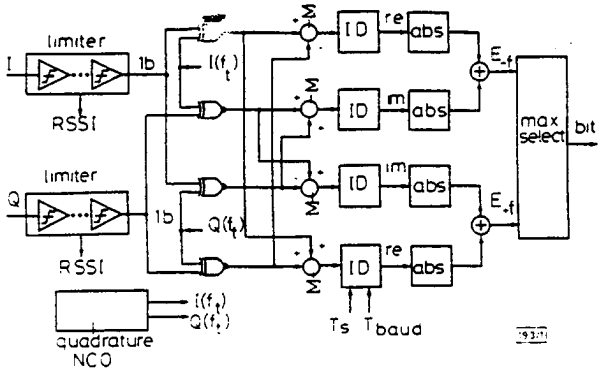


Fig. 1 *Multiplierless binary FSK correlation detector*

Implementation: In a sampled-data binary FSK correlation detector, the following magnitude is evaluated for the received signal $\tilde{s}$:

$$C(f_t) = \left| \frac{1}{N} \sum_{n=0}^{N-1} \tilde{s}(nT_s)\{\cos(2\pi f_t nT_s) + j\sin(2\pi f_t nT_s)\} \right| \quad (1)$$

where f_t is the tone frequency to be detected, and N is the oversampling ratio (T_{baud}/T_s). We have sought a simple implementation for this algorithm. Since FSK signalling requires only the frequency information and not the amplitude, the quadrature input signals are first hard-limited (Fig. 1). Before hard-limiting, a low-pass channel filter selects the baseband signal from neighbouring channels. Given the wide dynamic range (typically 80dB) of the radio channel, the baseband FSK detector must handle this range to avoid any gain control for RF signals. A CMOS limiting amplifier [6] capable of >80dB of dynamic range has been built to fulfill this requirement. However, limiting introduces odd harmonics of the original tones. These harmonics are aliased after sampling, potentially corrupting the orthogonality of binary FSK tones applied to the digital correlation detector. However, if N is an integer multiple of four ($4i$), it may be shown that distortion caused by harmonic aliasing is avoided. This guarantees that the generated harmonics are symmetric about half of the sampling frequency, retaining the Hermitian property of each I-Q signal. The detector must then only discriminate between signal energy at positive or negative frequency $\{+F_{tone}, -F_{tone}\}$.

The input signal must be correlated with the sine and cosine components for a quadrature correlator. Since the input signal is already hard-limited, the reference tone needs not be a pure sinusoid. In our approach, square waves with the proper tone frequency are used instead. Thus, an XNOR gate may be used as a 1 bit multiplier for signal correlation. Harmonics resulting from square waves can, after aliasing, downconvert undesired parts of the signal spectrum to baseband [7]. However, when the input signal is filtered and hard-limited, the spectrum at the harmonics is caused by the input signal. Thus, when the over-sampling ratio constraint is met, there is no extra degradation. The integrate-and-dump (ID) block is implemented with a simple accumulator, and its clocks are generated by a separate clock recovery loop [8]. As shown in eqn. 1, the correlation detector also requires a magnitude calculation unit. In our architecture, an absolute-value addition block replaces a conventional squaring multiplier. Thus, a truly multiplierless FSK detector is obtained with little performance degradation. The I-Q local tone generator is implemented with a 1 bit output numerically controlled oscillator (NCO), rather than a full-precision direct digital frequency synthesiser. The tone frequency to be detected is fully programmable by controlling the input control word of the quadrature NCO.

Measurements: Two basic requirements for binary FSK direct-conversion receivers are low power and high inherent dynamic range. All blocks shown in Fig. 1 have been implemented in 1μm CMOS, and consume 5mW of power with a supply voltage of 3V. Fig. 2 shows the measured dynamic range of the detector. Measurements were made with $F_{tone} = \pm 160$kHz, $F_{baud} = 160$kHz, $F_s = 10.24$MHz and $N = 64$. The inherent dynamic range of the detector is 82dBm at a BER of 10^{-3}. This result is higher than those reported for other FSK detectors [3, 9]. The minimum detectable signal power at a BER of 10^{-3} was measured to be −72dBm, which is mostly dominated by the input noise of the limiting amplifier. The upper limit, however, is set by the maximum output swing of the stage driving the limiter, which is 1 V peak-to-peak for the test case. For the detector alone, there is no inherent upper limit and the limiter input may swing to the power supply. Higher dynamic range may be achieved by dissipating more power in the hard-limiter stage. However, this is not necessary since 82dB of dynamic is sufficient for the radio channels encountered in most wireless applications.

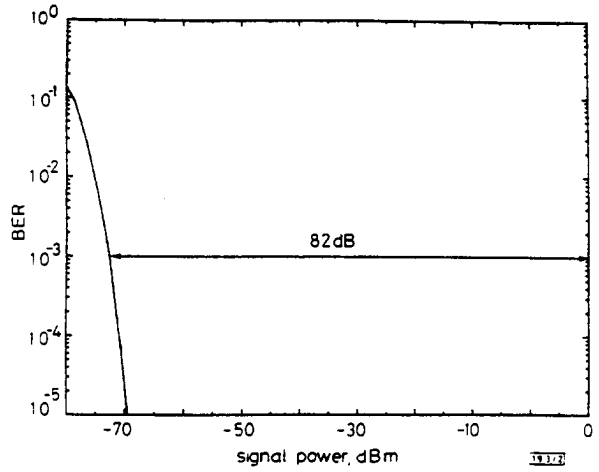


Fig. 2 *Measured dynamic range of binary FSK detector*

Conclusion: A multiplierless implementation of binary FSK correlation detection has been presented. Architectural simplifications applied to the design make it low power. With proper choice of oversampling ratio ($4i$), the proposed detector maintains the orthogonality of binary FSK tones without harmonic aliasing, and thus minimises performance degradation caused by 1 bit correlation. The measured dynamic range of the detector is 82dB at a BER of 10^{-3}. This FSK detector is suitable for monolithic integration into direct-conversion receivers used in wireless communications systems.

3 May 1995
Electronics Letters Online No: 19950717

J. Min, H.-C. Liu, A. Rofougaran, S. Khorram, H. Samueli and A.A. Abidi (*Integrated Circuits and Systems Laboratory, Electrical Engineering Department, University of California, Los Angeles, CA 90095-1594, USA*)

References

1 MIN, J., ROFOUGARAN, A., SAMUELI, H., and ABIDI, A.: 'An all-CMOS architecture for a low-power frequency-hopped 900 MHz spread spectrum transceiver'. Proc. IEEE Custom Integrated Circuits Conf., May 1994, pp. 379–382

2 SCHULTES, G., SCHOLTZ, A.L., BONEK, E., and VEITH, P.: 'A new incoherent direct conversion receiver'. Proc. 40th Vehicular Technology Conf., May 1990, pp. 668–674

3 KREUZGRUBER, P.: 'A class of binary FSK direct conversion receivers'. Proc. 40th Vehicular Technology Conf., June 1994, pp. 457–461

4 WILSON, J.F., YOUELL, R., RICHARDS, T.H., LUFF, G., and PILASKI, R.: 'A single-chip VHF and UHF receivers for radio paging', *IEEE J. Solid State Circuits*, 1991, **SC-26**, (12), pp. 1944–1950

5 PROAKIS, J.G.: 'Digital communications' (McGraw-Hill, New York, 1989), 2nd Edn.

6 KHORRAM, S., ROFOUGARAN, A., and ABIDI, A.A.: 'A CMOS limiting amplifier with signal-strength indicator'. To be presented at Symp. VLSI Circuits, Kyoto, Japan, June 1995

7 PROUDFOOT, A.D.: 'Simple multifrequency tone-detector', *Electron. Lett.*, 1972, **8**, pp. 524–525

8 LIU, H.-C., MIN, J., and SAMUELI, H.: 'A low-power baseband receiver IC for frequency-hopped spread spectrum applications'. Proc. IEEE Custom Integrated Circuits Conf., Santa Clara, CA, USA, May 1995

9 BONEK, E., SCHULTES, G., KREUZGRUBER, P., SIMBURGER, W., WEGER, P., LESLIE, T.C., POPP, J., KNAPP, H., and ROHRINGER, N.: 'Personal communications transceiver architectures for monolithic integration'. IEEE Int. Symp. Personal, Indoor, and Mobile Radio Communications, The Hague, The Netherlands, September 1994, pp. 363–368

Optimisation of modified Mueller and Müller algorithm

G.R. Danesfahani and T.G. Jeans

Indexing terms: Synchronisation, Phase shift keying

Optimisation of the modified Mueller and Müller (mM&M) timing error detection algorithm by cancelling its self-noise is presented. Simulation results show that not only is the new algorithm self-noise free, but there are also no symbol slips at medium to high SNR. Short acquisition time, operation on only one sample per symbol to estimate the timing error, and a low complexity are other features of the new algorithm.

Analysis: We are assuming a system whose output comprises nominally synchronised data symbols subject to additive white Gaussian noise, where the transmitter and the receiver use Nyquist root raised-cosine filters. With perfect carrier synchronisation, the sampled matched filter output is

$$p(k) = A_o \sum_{i=-\infty}^{\infty} a(i)g((k-i)T-\tau) + w(k) \qquad (1)$$

where $a(i)$ is the data symbol, $g(k)$ is the response of the receiver filter to the input pulse, A_o is the amplitude of the received signal, T is the symbol period, τ is the timing error, and $w(k)$ is additive white Gaussian noise. The mM&M algorithm [1] for the simple case of the BPSK modulation scheme is

$$\mu_1(k) = [\hat{a}(k-1) - \hat{a}(k+1)]p_r(k) \qquad (2)$$

where $\hat{a}(k)$ is the receiver's decision on $a(k)$. By substituting eqns. 1 and 2 for $p(k)$, the mM&M algorithm can be expanded into

$$\begin{aligned}\mu_1(k) = {} & A_o[\hat{a}(k-1)-\hat{a}(k+1)] \\ & \times \{a(k+1)g(-T-\tau)+a(k-1)g(T-\tau)\} \\ & + A_o[\hat{a}(k-1)-\hat{a}(k+1)]a(k)g(-\tau) \\ & + A_o[\hat{a}(k-1)-\hat{a}(k+1)] \\ & \times \sum_{i=\infty, i\neq k, i\neq k\pm 1}^{\infty} a(i)g((k-i)T-\tau) + w(k)\end{aligned} \qquad (3)$$

The terms on the first line contribute to the error voltage, which is proportional to the timing error τ. The term on the second line is the self-noise which will not disappear, even when $\tau = 0$. With a Nyquist pulse shape, the summation term on the last line vanishes during tracking. To cancel the self-noise, the following is added to the self-noise:

$$[\hat{a}(k)a(k+1) - \hat{a}(k)a(k-1)]g(-\tau) \qquad (4)$$

At high SNR, the receiver's decisions are correct and, therefore, the self-noise entirely vanishes. From eqn. 1 it can be deduced that adding eqn. 4 to eqn. 3 means adding the following to the original algorithm:

$$\mu_2(k) = \hat{a}(k)[p_r(k+1) - p_r(k-1)] \qquad (5)$$

The mM&M synchroniser and the algorithm in eqn. 5 generate sample outputs proportional to timing error, but the jitter components are in anti-phase so that if

$$\tau = 0.5[\mu_1(k) + \mu_2(k)] \qquad (6)$$

the jitter is minimised. To generalise eqn. 6 to include QPSK and OQPSK modulation schemes, the mM&M algorithm is

$$\mu_\tau(k) = \mathbf{R}\{[\hat{c}(k-1) - \hat{c}(k+1)]p^*(k)\} \qquad (7)$$

where $\hat{c}(k) = \hat{a}(k) + j\hat{b}(k)$ is the receiver's decision about the real and imaginary components of data, $p^*(k) = p_i(k) - jp_q(k)$ is the complex conjugate of the real and imaginary components of the sampled matched filter output, and $\mathbf{R}\{x\}$ is the real part of the complex value x. By going through eqns. 2 - 7, the generalisation of the optimised algorithm becomes

$$\mu(k) = \mathbf{R}\{[\hat{c}(k)-\hat{c}(k-2)]p^*(k-1)+\hat{c}^*(k-1)[p(k)-p(k-2)]\} \qquad (8)$$

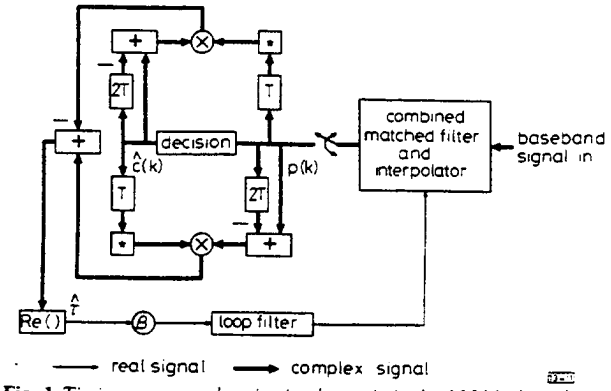


Fig. 1 *Timing error synchronisation by optimised mM&M algorithm*

The block diagram of the optimised mM&M algorithm for the QPSK modulation scheme is shown in Fig. 1. The output of the combined matched filter and interpolator [2] is sampled at symbol rate, $1/T$. Therefore, the operation of the optimised algorithm requires one sample per symbol. To detect the timing error, $2N_i+4$ real multiplications and $2N_i+5$ real additions are performed per timing error estimate $\hat{\tau}$, where N_i is the number of filter coefficients in the interpolator FIR subfilters.

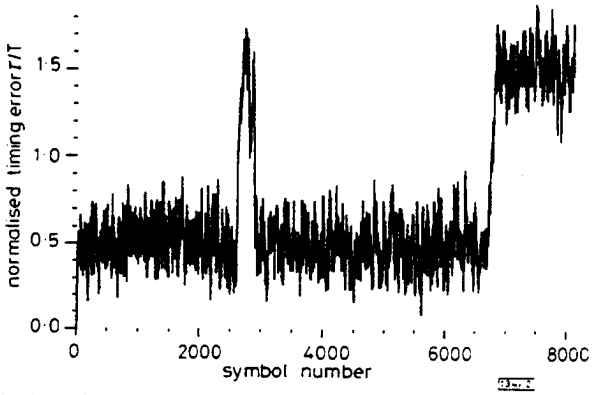


Fig. 2 *Performance of mM&M algorithm*

Simulation results: A QPSK modem was simulated using COSSAP, and the optimised algorithm was incorporated in the demodulator. Fig. 2 shows the normalised timing error of the data output for an initial timing error of half a symbol period and loop gain factor $\beta = 0.18$ using the standard mM&M algorithm. These results assume perfect carrier recovery, and high SNR at the input of the modem. It can be seen that although the timing loop has a fast acquisition, there is a large tracking jitter due to self-noise. In addition, there are symbol slips which degrade the error performance of the receiver.

Fig. 3*a* shows the performance under the same conditions when the optimised algorithm is used. The timing jitter caused by self-noise has been greatly reduced, and the symbol slips have been eliminated. The fast acquisition characteristics of the mM&M algorithm have not been affected, and Fig. 3*b* shows the acquisition performance for an initial timing error of half a symbol period.

The simulations have been repeated for signals subject to additive white Gaussian noise with SNR from 0 to 15dB, and results have shown consistently better results than those using the original algorithm [3]. These further results are being prepared for publication.

Low Power Radio-Frequency ICs for Portable Communications

A.A. Abidi

Proceedings of the IEEE
vol. 83, no. 4, pp. 544-569
April 1995

Low-Power Radio-Frequency IC's for Portable Communications

ASAD A. ABIDI, MEMBER, IEEE

Invited Paper

The contributions of integrated circuits to the RF front-end of wireless receivers and transmitters operating in broadcast and personal communications bands are surveyed. It is seen from this that when IC's enable a rethinking of the RF architecture, the wireless device can sometimes become significantly smaller, and consume much less power. Examples are taken from FM broadcast receivers, pagers, and cellular telephone handsets.

Many semiconductor technologies are competing today to supply RF-IC's to cellular telephones. The various design styles and levels of integration are compared, with the conclusion that single-chip silicon transceivers, combined with architectures which substantially reduce off-chip passive components, will likely dominate digital cellular telephones in the near future.

The survey also projects future trends for IC's for miniature spread-spectrum transceivers offering robust operation in the crowded spectrum. With sophistication in baseband digital signal processing, its increasing interaction with the RF sections, and with increasing experience in simplified radio architectures, all-CMOS radios appear promising in the 900 MHz to 2 GHz bands. A specific CMOS spread-spectrum transceiver project underway at the author's institution is discussed by way of example.

I. INTRODUCTION

The portable revolution is upon us today. It promises to empower individuals throughout the world by giving them low-cost access to information wher ever they may be, thus allowing them to make informed decisions and to be more productive in business and at home, without necessarily being tied down to a physical location. It is expected that in the near future, individuals will be equipped with capabilities of local computing and of communications enabling them to perform almost all the tasks that today require the equipment on the office desktop: the telephone, the computer, its connection to a ubiquitous wired network, the fax machine, and so on [1].

Manuscript received November 21, 1994; revised January 13, 1995. This work was supported in part by the US Advanced Research Projects Agency, and in part by a consortium of semiconductor companies under the State of California MICRO Program.

The author is with the Electrical Engineering Department, University of California, Los Angeles, CA 90024-1594 USA.

IEEE Log Number 9409344.

The portable revolution has been many years in the making. The personal broadcast radio receiver and cassette player, as pioneered by Sony, was a runaway global success. The user could construct a private audio environment anytime and anywhere, using a device so small and light that its presence was easily forgotten. Portable computers brought about the next wave of change. The personal computer has had such a large impact on the broad working habits of individuals that without it they are lost. The luggable computer has evolved in a matter of one decade into the portable, the notebook, and the subnotebook. With the establishment of a wide-area network of radio paging transmitters, the personal radio pager has also became very popular since the mid-1980's. Today, the pager network spans the entire continental US and many other parts of the world, enabling the user to receive alphanumeric electronic mail messages. Two-way paging is actively under development. The cellular telephone became widely available shortly thereafter. The user could hook into the international switched-telephone network through the nearest cellular base station with a portable transceiver, which too has scaled down remarkably in size until its weight and volume is the smallest practical [2].

There are many competing visions of how these various portable devices and services will evolve and integrate in the next few years. These are covered extensively in the popular press, and in numerous keynote speeches in technical meetings. A common theme is that users will want a multimedia terminal, capable of wireless access to a global network which can transport communications, images, and databases to the user in an on-demand, interactive fashion [3], [4]. Such a terminal will have capabilities of computing, image acquisition and display, and obviously, of communications. It will likely derive as a hybrid of the various portable technologies available today.

What obstacles must be overcome to realize this vision? There include the very highly integrated electronics, effective displays, and a philosophy of design based on low power dissipation to prolong the battery life of the portable

device. Sometimes single-battery operation will impose the additional constraint of operation at low voltage, as low as 1 V, which will require entirely new ways of doing electronic circuit design.

In the past few years, most designers of mass-market digital IC's have been preoccupied with low-power operation [5]. Principles such as operating CMOS logic at the lowest possible supply voltage have become widely known, and power-down modes, gear-shifting of operating clock frequencies, pipelining and parallelism, subthreshold operation, and other such methods which were once the province of specialized areas such as electronic wristwatches and implantable biomedical devices are becoming commonplace [6]. Studies into the fundamental thermodynamic limits to the energy required for computation are being initiated or revived. There is good reason to believe that all this activity will lead to significant improvements in the conventional circuit and system design styles for digital signal processing and computation.

How will this activity affect the communications aspects of the portable device? What similar principles to low-power digital design are there for energy-efficient wireless communications transceivers? These questions do not have simple answers. Low-power communication systems will result from use of the correct architectures, a sensible partition between analog and digital signal processing, low-power circuit techniques everywhere, and a judicious division between active and passive components. There is still not a widely known, integrated vision on this subject. Furthermore, there remains a gap between the IC design used in the portable applications described above, and the new designs that will be required over the next few years for advanced portable communicators. Consumers are demanding a great deal more functionality and performance, which is stressing present-day technology to its limits, and wireless communicator design itself is in transition from the classic analog modulation techniques used over the past 50–70 years to more sophisticated methods using digital signalling formats and signal-processing methods in transceivers.

This paper summarizes the key developments in the discipline so far, and from them forecasts wireless IC design trends in the near future.

II. Key Signal-Processing Issues in Wireless Transceivers

Were it not for the advent of the portable communications revolution, radio technique would almost certainly have become a lost art. The first edition of the last definitive textbooks on the subject dates to 1943 [7]. Whereas once radio engineering was synonymous with electronics [8], few university electronics curricula today offer a course on radio communications circuits. Only a few modern textbooks on radio design have been written in the past 25 years [9]–[13]. Radio communication methods, at least for nonmilitary applications, have remained relatively unchanged since World War II, and the evolutionary improvements in consumer equipment mainly owes to the use of high-frequency discrete transistors [14], smaller passive components, and building-block IC's which improve the long-term reliability and manufacturability of radio and TV receivers. The major impact of IC technology in these consumer items has probably been at baseband, in adding more user features. In contrast, the front-end radio architectures have evolved almost not at all in the past 40 or so years. For instance, IC's have contributed digital volume-control, digital frequency-tuning, features to alleviate manual effort on the part of the user, but the RF and IF sections still contain discrete and passive components in rather conventional architectures.

Why is this? It is partly because radio frequencies were too high for the low-cost IC technologies traditionally used in the consumer electronics industry. It is also because advances in component packaging alone have led to rapid downscaling in the size of consumer devices, often obviating the need to rethink the electronics. As a result, only a few individuals in a handful of institutions worldwide have concerned themselves with thinking about these problems. Today, as conventional solutions no longer suffice for the future wireless communications devices, there is a rekindling of interest in this subject, which has led to much rediscovery and some invention. Baseband IC designers are now attempting to apply familiar techniques to wireless, while microwave IC designers are exploring what to them are low-frequency commercial opportunities for their technologies.

To set the stage for further discussion, some of the unique problems of radio receivers and transmitters are first described. Unlike familiar wireline communications, the wireless environment accommodates essentially an unlimited number of users sharing different parts of the spectrum, and very strong signals coexist next to the very weak. The radio receiver must be able to select the signal of interest, while rejecting all others. It must do so using less than perfect active and passive components. There are two important problems in the receiver: *image-rejection* and *dynamic range*. Image-rejection relates to the receiver's ability to select the desired signal from the array of signals occupying the spectrum. Ideally, it might do so with a tunable bandpass filter, whose center frequency could be positioned at will in the RF, and whose passband was one channel wide. A filter with this small a fractional passband does not exist. Instead, a practical RF bandpass filter, which may or may not be tunable, will preselect an array of radio channels including the one of interest (Fig. 1). The other preselected channels are then removed at a lower intermediate-frequency (IF), by translating them in frequency with a downconversion mixer, and centering the desired channel within a bandpass filter at IF. The other mixer input is a frequency-tunable local oscillator (LO), offset by IF from the desired channel. As the preselected band after downconversion will very likely occupy an interval greater than (0, IF) on the frequency-axis, the IF bandpass filter will select both the desired channel, and another *image* channel the mixer has translated to –IF. The subsequent detector circuit will be unable to distinguish

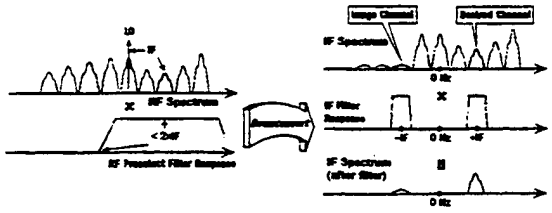


Fig. 1. The image-reject problem in radio receivers. The RF preselect filter passband must be determined with prior knowledge of the IF. The preselect filter is responsible for image-channel suppression before downconversion.

between the desired and the image channels, and therefore its output will be the result of the superposition of both. However, if the stopband of the preselect filter lies less than $2 \times$ IF away from the desired channel it will attenuate the image, so only the desired channel will contribute energy at IF. The receiver designer first studies the available filter technologies, and then chooses an appropriate IF which yields an acceptable image suppression. A high IF relaxes the prefilter passband specification, but it also means that the downconverted signal requires high-frequency amplifiers, which are usually power-inefficient. Further, the IF filter requires a smaller fractional passband. In such cases, following image rejection at this high IF, the channel may be selected after downconversion to a second, lower IF. In such a double superheterodyne, or dual-conversion receiver, the first IF may actually lie at a higher frequency than the incoming RF to make image rejection easier.

The noise-level and nonlinearity in the RF amplifier and first mixer usually set the receiver dynamic range. Consider reception of a weak channel surrounded by large undesired channels in the preselection band (Fig. 2(a)). First, the input-referred noise of the receiver directly adds to the sought signal, corrupting its signal-to-noise ratio (SNR). Second, the large adjacent channels will experience the nonlinearities in the RF amplifier and mixer, and some of the products of the ensuing intermodulation distortion may overlap the desired channel. The IF filter cannot reject these unwanted products, which, like noise, will degrade the received SNR. As the receiver frequency response is normally bandpass, its nonlinearity is measured by applying two tones of equal amplitude closely spaced in frequency (f_1 and f_2) at it input, and measuring the rise in the third-order intermodulation products (at $2f_1-f_2$ and $2f_2-f_1$) with input level (Fig. 2(b)). All other intermodulation tones usually lie outside the receiver passband. On a logarithmic plot, the third-order intermodulation level rises at a slope of 3 relative to the fundamental tone at the output. The two lines intersect at a point called the input-referred 3rd-*order intercept* (IP3). The intercept point is usually extrapolated from measurements at low levels, because the receiver front-end will saturate at large inputs. The 1-dB *compression point*, the input level which causes the receiver gain to drop by 1-dB relative to the small-signal gain, specifies the onset of saturation. The input-referred noise-level may be included in this plot to define a spurious-free dynamic range (SFDR), although this is rarely used in radio

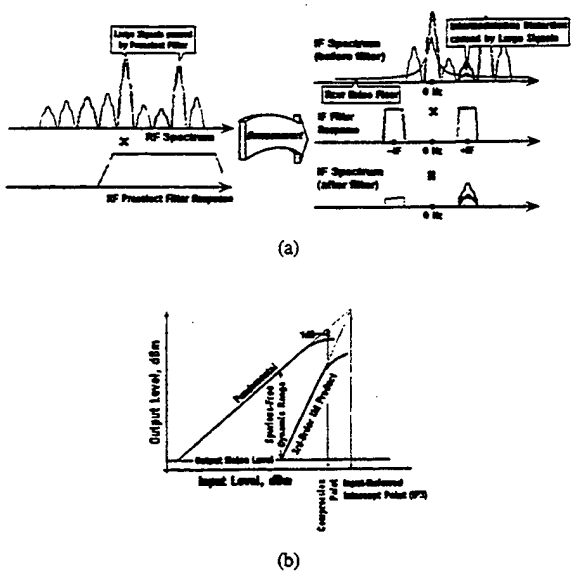


Fig. 2. The dynamic range problem in the radio receiver. (a) Adjacent large signals may create intermodulation products superimposed on the desired channel. Receiver noise floor is fundamental limit to sensitivity. Receiver dynamic range is specified (b) in terms of extrapolated 3rd-order intercept point, and noise level.

specification. Usually the input noise-figure (NF) and the input-referred intercept point (IP3) are separately specified.

Similar specifications apply to the transmitter, which operates at much larger signals. Suppose, as is almost always the case, that the transmitter is required to emit a single-sideband, suppressed-carrier output. Owing to circuit imperfections, it may also emit small amounts of the carrier and the unwanted sideband, which typically lie in the passband of the subsequent RF filter (Fig. 3). These unwanted emissions may become interferers for adjacent channels. Nonlinearities in the power amplifier may also produce emissions of intermodulation products at other frequencies. Phase-noise in the local oscillator responsible for upconverting to RF will convert to noise added to the signal, and the amplitude of this noise increases with the transmitted signal. This noise could possibly overwhelm nearby weak channels. Transmitter performance is usually specified in terms of the relative levels of unwanted signals to the desired signal, and in terms of absolute spectral density of output noise at maximum output power.

To understand the rationale underlying receiver architecture, let us use as an example the familiar broadcast FM receiver. The architecture to be described is the same that Armstrong, the inventor of FM and the superheterodyne, had originally proposed for FM reception. The desired channel consists of a carrier in the 88–108 MHz band, modulated by up to ±75 kHz. Neighbouring channels are spaced apart by 200 kHz. The receiver must select the desired channel while rejecting nearby channels, and it must be sensitive to a signal of a few tens of microvolts induced on the antenna. A simple FM antenna is wideband, and will pick up signals well outside the broadcast FM band. The low-noise bandpass amplifier in the front-end may at best mildly attenuate the out-of-band signals—the

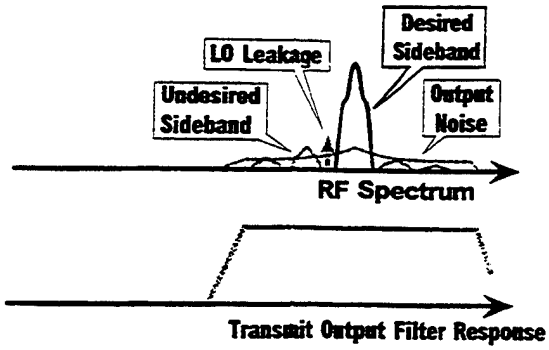


Fig. 3. Transmitter imperfections (such as mismatches in a quadrature upconverter) result in the appearance of spurii in the emitted spectrum (referred to as "spectral regrowth"). The spurii may not be removed by the output filter, and may superimpose on adjacent channels. Emitted noise might overwhelm weak adjacent channels.

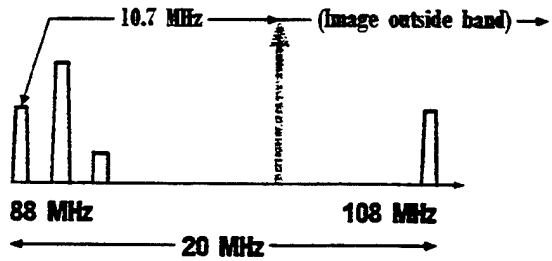


Fig. 4. The 10.7 MHz intermediate frequency conventionally used in FM receivers is the lowest frequency which will guarantee that the image lies outside the FM broadcast band. The FM demodulator will reject a (non-FM) image channel entering the receiver. 10.7 MHz passive bandpass filters in the IF strip select the desired channel.

actual channel selection must be done elsewhere. The RF amplifier uses an inductive load to resonate with the transistor and tuning capacitances, thereby transforming an inherently low pass characteristic to bandpass centered on the frequencies of interest. The transistor f_{max} limits the highest frequency at which such a tuned amplifier can still provide a gain greater than unity. This figure-of-merit is familiar to microwave circuit designers and to device designers, whereas baseband IC designers deal more often with transistor f_T, the capacitance-limited unity current-gain frequency. In bipolar IC processes not optimized for small-signal high-frequency use, f_{max} is comparable to f_T, whereas in the best RF processes it may be twice f_T [15], [16].

It is impossible to select the desired channel at RF, because no tunable filters exist with the required fractional bandwidth of 0.15% at 100 MHz. Therefore, following sufficient amplification at RF to overcome the noise-level of the following circuits, the signal is mixed down by a variable-frequency local oscillator to a lower IF. Furthermore, a filter to select the desired channel at IF will have a fixed center frequency, and the fractional bandwidth in the passband will be more reasonable. From this perspective, it is desirable to use as low an IF as possible. However, image-rejection poses yet another constraint on choice of IF. Conventional broadcast FM receivers use an IF of 10.7 MHz as a compromise. This IF guarantees that the image always lies outside the FM band (Fig. 4). It is unlikely, however, that the preselect filter can suppress this image, which will therefore either add noise or AM to the desired signal. However, the subsequent FM detector is inherently insensitive to both these forms of impairment. In this way, a medium-valued IF is made possible by exploiting properties of the detector, and thereby the fractional passband specification for the preselect filter is relaxed. The first mixer downconverts the entire FM band, with the desired channel centered at 10.7 MHz. A varactor-tuned Colpitts oscillator may be used as the first local oscillator. Prior to detection, a cascade of identical, fixed-frequency ceramic bandpass filters, each with a 200 kHz passband centred at 10.7 MHz, passes the desired channel while rejecting neighbouring channels [17]. In a high-quality receiver, the RF amplifier may be a discrete GaAs MESFET with a tuned load ganged to the LO tuning element, and another ganged tuned circuit may couple the antenna signal into the receiver [17].

To transplant this style of discrete radio-circuit design to IC's, one would have to implement LC tuned circuits and filters on silicon. One can indirectly surmise these concerns in the Motorola series of IC design textbooks from the 1960's, which discuss loss in spiral metal inductors fabricated on silicon substrates [18], as well as issues relating to simulated inductors for active filters at IF for radios, made with gyrators and capacitors [19]. However, on-chip spiral inductors of useful values were found to suffer excessive capacitance to the substrate, which lowered their self-resonant frequency to the point that they were not usable beyond the VHF band. It gradually became part of the collective consciousness of IC designers that on-chip tuned circuits are generally impractical.

When useful tuned amplifiers did appear on monolithic integrated circuits, it was not for the VHF to UHF range of relevance to consumer applications. Instead, it was military applications at much higher frequencies which drove the development of monolithic microwave integrated circuits (MMIC's) during the 1980's. MMIC's typically use MESFET's as the active device on semi-insulating GaAs substrates. This technology has enabled miniature radar, remote sensing, and communications at frequencies up to tens of GHz. The on-chip wavelengths are so small that monolithic distributed circuits may be built. MMIC's take advantage of the semi-insulating substrate in two important ways. Transistors on these substrates have lower parasitic capacitance, which means that they amplify to higher frequencies. It is also possible to build low-capacitance interconnect with airbridge structures, and high-frequency passive components such as spiral inductors required for narrowband tuned circuits. Thus on MMIC's, the board-level design styles used hitherto by radio- and microwave-engineers could be miniaturized. However, over its many years of existence, GaAs MMIC technology has not had the major impact on consumer electronics that its adherents had hoped for. Makers of consumer electronics favor silicon IC technology wherever feasible because of its low-cost, high

yields, and the relative ease of mixing analog and digital circuits on a large scale.

It is anticipated that by the year 2000 about 300 million portable consumer wireless devices will be in use [20]. What IC technologies will enable the RF front-end of these devices? Do miniaturization and long battery-life call for architectural innovations in transceivers? What new circuit design styles will evolve in response? There is much curiosity and speculation on these matters, yet little is generally known about RF-IC design, or on the possible impact of large-scale integration and power-reduction strategies in the front-end of wireless transceivers. This paper presents a brief survey of the use of IC technology in wireless receivers and transmitters since the 1970's to date, and from this projects some future trends. RF-IC's are roughly defined as integrated circuits operating in the band of frequencies from 400 MHz to 2500 MHz, which covers most consumer wireless communication devices. As opposed to MMIC's, which were almost exclusively fabricated on III-V compound semiconductor substrates at small-scales of integration, mature silicon technologies will play a large, if not the dominant, role in RF-IC fabrication. It is the author's belief that in response to pressing demands for ubiquitous wireless access, both the underlying semiconductor technology and the design styles will rapidly evolve to realize the single-chip "VLSI radio" in the not too distant future.

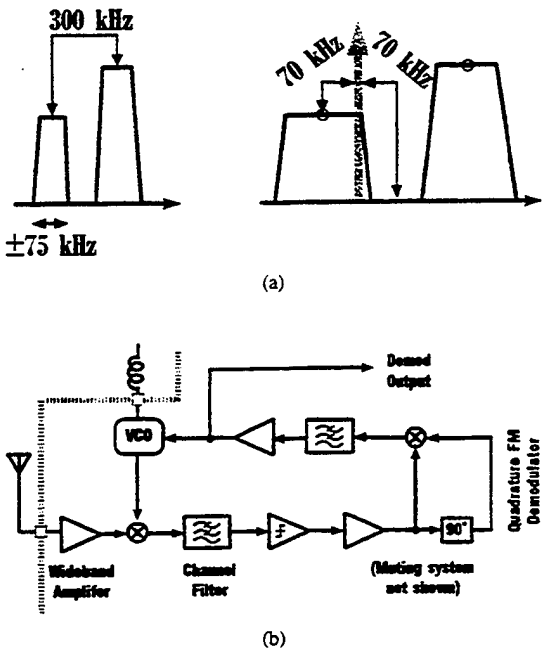


Fig. 5. (a) An alternative choice of IF in the FM band, which places the image in the gap between adjacent channels. The IF strip, including the channel filters, now operate at a 70 kHz frequency. (b) An FM receiver using a 75 kHz IF. The channel filter is an active-*RC* implementation on-chip. A frequency feedback loop compresses the incoming frequency swing. Except for a tuning inductor, no high-frequency off-chip components are required.

III. IC's in Broadcast Radio Receivers

The two-way wrist radio has fascinated the popular imagination since its introduction in the popular American cartoon strip, *Dick Tracy*, in 1946 [21]. This is the ultimately unobtrusive piece of consumer electronics. Let us now see how feasible it is to build an FM receiver of this size with microelectronics technology. For the average user to accept such a radio, its selectivity and sensitivity must be comparable to that of tabletop models. If in the FM receiver described in the previous section all the transistors in the receiver electronics were to be integrated on to one silicon chip, the radio would still need a considerable number of off-chip tunable inductors and ceramic filters, and in spite of state-of-the-art miniature packaging, the components could not plausibly all fit into a wristwatch. Neither would the power dissipation be commensurate with the life of a wristwatch battery.

The first generation of silicon bipolar IC's developed in the late 1970's for the IF and baseband portions of broadcast receivers more or less contained the transistors of conventional receivers assembled on to one or more IC's [22]–[24]. However, integration did afford freedom to use transistor-rich circuits for higher performance. Circuit techniques such as double-balanced mixers using the Gilbert analog multiplier, phase-locked loops as FM demodulators, and balanced on-chip signal paths to attain greater immunity to pickup and common-mode noise became widely used as a result. By eliminating many of the coupling coils and other noncritical discrete components found in older radio circuits, IC's contributed to lowering the cost of assembling and aligning the final product. In the RF section, though, the receivers still used the conventional 10.7 MHz IF superheterodyne architecture implemented with shielded discrete-component circuits.

In the early 1980's, Kasperkovitz at Philips [25] made the first significant explorations into alternative architectures for highly integrated radio receivers. He realized that to reduce receiver size and power dissipation, it was very important to eliminate the many off-chip passive components. If certain passive inductors, capacitors, and resistors could not be eliminated, they could at least be packaged in surface-mount outlines for very small size. However, neither the volume of the IF ceramic filters could be readily scaled down, nor could their characteristic impedance be scaled much above 50Ω. Each filter requires an on-chip analog driver of comparable impedance. Alternatively, the filter may be realized on-chip as an active bandpass circuit of sufficient selectivity and dynamic range. Although a gyrator-capacitor based active filter is possible in principle, small phase-shifts in the gyrator transistors at the 10.7 MHz IF can seriously upset the filter passband shape. Active resonators are also known to suffer from a larger internal noise level than their passive counterparts, and this discrepancy worsens with increasing pole-Q and pole frequency [26]. Kasperkovitz solved the problem with an *architectural innovation*, by dramatically lowering the IF from 10.7 MHz to 70 kHz. This makes it a great deal easier to implement an IF active channel-select filter, which

now need only be lowpass. Further, at a given dynamic range, the power dissipation in an active filter also scales down with the IF [26]. The low IF eliminates the off-chip channel-select filter and reduces power dissipation, both very desirable properties. But what of the image frequency, the principal reason for the choice of 10.7 MHz?

At a 70 kHz IF, the image frequency lies half-way to the adjacent FM channel (Fig. 5(a)). The image therefore is the inter-channel noise in the FM band. As an RF preselect filter cannot possibly reject an image this close to the desired signal, it will pass unattenuated to worsen the received signal-to-noise ratio (SNR) by 3-dB. Another consequence of this choice of IF is that after the first downconversion, an instantaneous frequency deviation in the received FM signal of more than 70 kHz will alias around dc to produce distortion. This is avoided by compressing the frequency deviation to ± 15 kHz with a negative feedback frequency-locked loop prior to downconversion (Fig. 5(b)). The FM mono/stereo radio [27]–[29] requires, in addition to the single-chip receiver, only 15 small capacitors and two inductors, and this collection of parts readily fits inside a wristwatch. A miniature earphone is plugged into the watch, and the earphone lead serves as the antenna. The radio when active drains 8 mA from a 4.5 V supply.

Sony, one of the world's leading makers of miniature radios, has also recently modified its integrated FM radio-receiver architecture from the conventional 10.7 MHz IF [30] to low-IF [31]. In the new architecture, a high first-IF of 30 MHz is used, so any out-of-FM band image falls in the stopband of a fixed 80 to 110 MHz bandpass preselect SAW filter after the antenna. A wideband IF amplifier boosts the received signal level of *all* FM channels *without* any filtering—amplification at 30 MHz is not a problem on this modern silicon bipolar IC process. Further, at a 30 MHz IF the entire broadcast FM band falls to one side of the LO frequency, which means that any channel in the FM band may be selected after the first downconversion. Following this, another mixer converts to a low second-IF of 150 kHz, and thereafter an on-chip 9th-order active-*RC* lowpass filter rejects adjacent channels. The low second-IF, however, will pass an image FM channel as well as the desired channel, and as there is no filtering at all at the first IF, an *image-reject mixer* is used for the second downconversion (Fig. 6). The variable-frequency first LO tunes the desired channel, while the second LO, which must produce quadrature outputs for the image-reject mixer, is at a fixed frequency.

The image-reject mixer provides a trigonometric solution to a difficult filtering problem [32]. The desired channel and its image are frequency-converted into two paths by mixers driven by quadrature phases of an LO. The mixer outputs are then phase-shifted 90° with respect to one another. The sum of these two signals will select the desired channel and suppress the image, while, *vice-versa*, the difference will select the image. The extent of image suppression depends on the gain matching of the two paths, and on the phase-accuracy of the LO quadrature outputs. For these reasons, this concept has only become practical with IC technology,

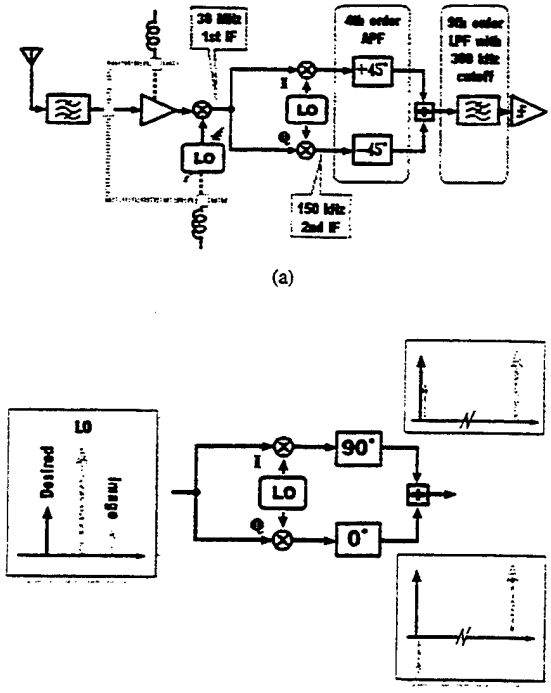


Fig. 6. (a) An alternative architecture for a single-chip FM receiver, with a 30 MHz IF chosen for strong image-rejection, and to translate the all channels in the FM band to IF on to one side of dc on the frequency-axis. The image-rejection mixer at 150 kHz selects the desired channel, while rejecting the undesired one 300 kHz away. Aside from a noncritical RF preselect filter, the remaining filters are active on-chip, including the phase-shifts in the two arms of the mixer. (b) The image-rejection downconversion mixer. The image and desired tones are at positive- and negative-frequency offsets from the local oscillator, and are discriminated in the mixer by a relative inversion of polarities in the two arms.

where the two paths are well matched on-chip and track each other over temperature. Image suppression on the Sony chip is limited to about 40–45 dB by residual gain mismatch in the two paths. An allpass active *RC-CR* filter produces 90° phase-shifted versions of the downconverted input. This receiver drains about 15 mA from voltages as low as 0.9 V in either FM or AM mode. All the necessary transistors are integrated on-chip—the RF amplifier portion, however, uses an off-chip load inductor and the local oscillator needs an off-chip *LC* tuned circuit.

The circuit techniques which enable sub-1 V operation are also interesting (Fig. 7). The antenna signal drives the emitter of a common-base NPN, which forms the tail of a differential pair. The input resistance of the common-base stage matches the antenna impedance. The signal develops at one inductively loaded collector of the differential pair, while the other collector dumps a fraction of the signal current into the supply in response to an AGC differential control voltage. Following amplification by a resistively loaded differential pair in cascade, the balanced RF signal is level-shifted into the first mixer, a simplified double-balanced Gilbert-cell with resistors instead of current sources in the tails.

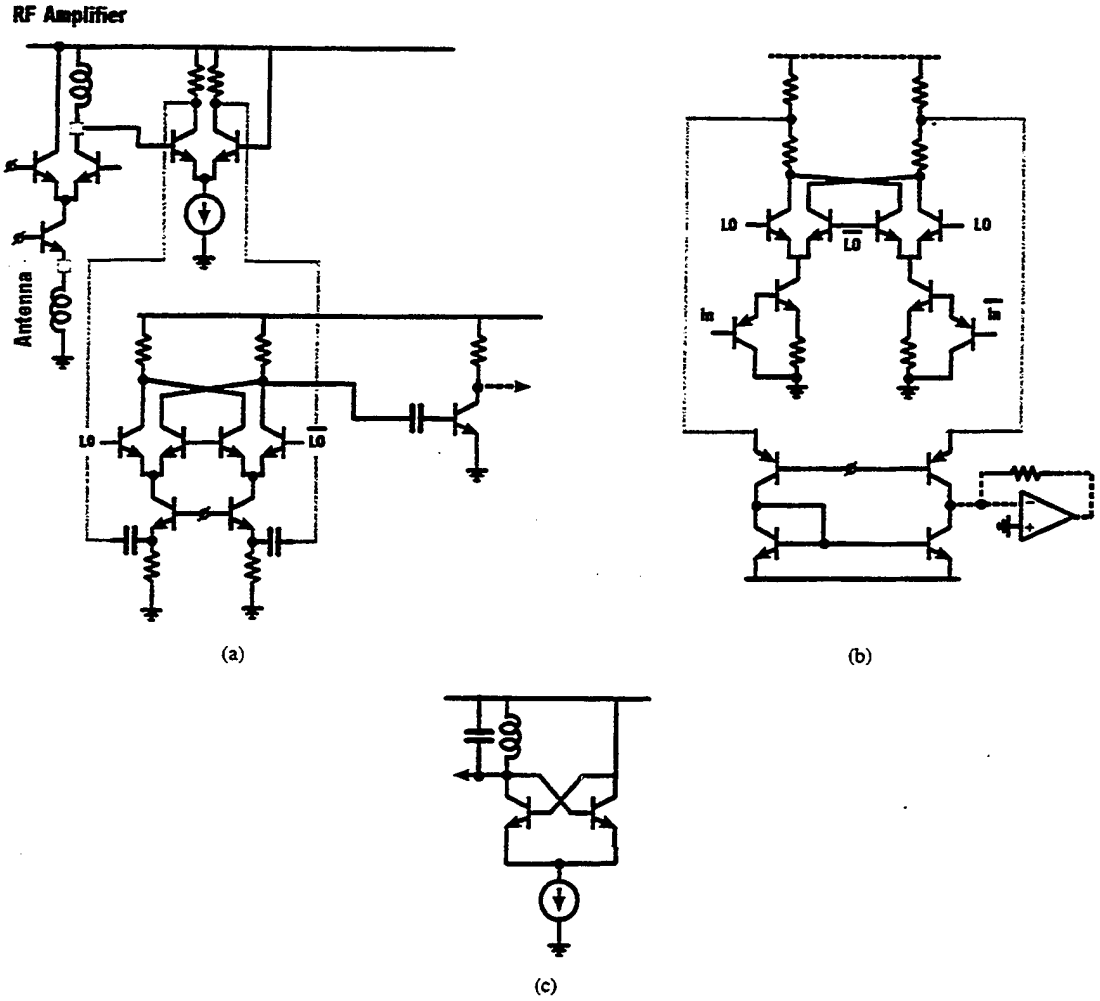


Fig. 7. Low-voltage circuits capable of operation with a 0.9-V supply. (a) RF amplifier. Note how the antenna carries the RF amplifier bias current, and how the on-chip capacitors level-shift the RF signal path. (b) IF amplifier. As the signal is downconverted in frequency, active level-shift and op amp circuits appear. (c) The cross coupled differential pair negative resistance is popular for LC oscillators.

As inductors do not drop a dc voltage, they make convenient loads for stacked transistor circuits operating at a low supply voltage. The signal on the free end of the inductor will swing above the supply. This is the case in the local oscillator, which implements a negative resistance of $-1/g_m$ with a cross coupled differential pair, causing an LC tuned circuit across it to oscillate. The oscillation amplitude is limited by the differential pair nonlinearity to a small multiple of kT/q.

As the signal propagating through the receiver downconverts in frequency, low-frequency circuit techniques and devices with a lower f_T are used. For instance, the 30 MHz IF signal couples into the second mixer through PNP emitter followers. At the 150 kHz IF, common-base PNP's and current mirrors level-shift down the signal, which is further amplified in an op amp.

This receiver is a significant example of how architectural rethinking combined with appropriate IC design styles has resulted in a very different solution to the well established broadcast FM receiver.

IV. IC's in Wireless Paging Receivers

Miniature wireless communicators to page people on the move were first developed in the late 1950's. The Bell System's Bellboy™ paging receiver [33] anticipated many of the concepts underlying today's pagers. The system operated at a 150 MHz RF, addressing a receiver by frequency-modulating the carrier with a unique set of three tones, which the intended receiver recognized at baseband by the simultaneous response of three passive reed resonators. The superheterodyne receiver operated at 4 V using a total of only ten transistors [34], a notable early example of low-power and low-voltage circuit design. The low IF of 6 kHz meant that simple, capacitively coupled 10 kHz lowpass (rather than bandpass) filters could select the desired channel. Also, all low frequency amplifiers after the RF section used transistors biased at small currents. The two stages in the cascode RF amplifier were inductively coupled to share, or reuse, the same bias current, a power-saving method found even in todays MMIC's.

veloped similar zero-IF bipolar integrated front-ends [45], [46]. A notable feature of the chips is that much of the die area is taken up by the capacitors for ac coupling the signal path and for the on-chip lowpass filters. Most pager IC's operate at supplies of 2 V to as low as 1V, and are implemented in silicon bipolar technology. In addition, full-featured pagers require 20 000 gate-equivalent digital IC's for the user interface [47], low-voltage EEPROMs for customization and software, and capability to drive a liquid-crystal display. The basic paging receivers can fit within a wristwatch [48]. The direct-conversion FSK digital paging receiver concept has also been successfully used at very low carrier frequencies (100's of kHz) with much lower data rates in implanted devices for biomedical applications [49].

V. IC's in Cellular Telephone Transceivers

Mobile and handheld cellular telephones are the first widespread two-way radios for consumer use. They were preceded by cordless telephones for local-area use. These wireless telephones must meet stringent demands for low weight and volume, long battery life, low cost, and reliable network access to be successful with consumers. In contrast, walkie-talkie transceivers were always aimed at specialized markets, and did not face these pressures for miniaturization. The average consumer, for instance, will not voluntarily accept a transceiver of the size and weight that policemen or soldiers carry as part of their outfit. Further, to support large numbers of users in a crowded radio spectrum, wireless telephones use more internal signal processing than other common transceivers, and must be capable of connecting to the public switched-telephone network [50]. Features such as digitally selected channels, direct-sequence spread spectrum, and diversity-selection are now becoming common. The transceivers perforce must use highly integrated, low-power electronics. Thus it may be said that with the advent of the modern cellular telephone the conventions of wireless design are being reexamined, and sometimes rewritten.

The first generation of cellular telephones carried voice signals by analog frequency modulation of a carrier. In the US AMPS system, for example, the handset receives at a carrier selected from the 869–894 MHz band, while it transmits on a carrier in the 824–849 MHz band. These 25 MHz wide bands are separated by 45 MHz between the uplink and downlink, enabling the user to talk and listen at the same time much as on the wired telephone (users are not too fond of the "over, over-and-out" protocol). An antenna duplexer suppresses coupling from the transmitter into the sensitive receiver, acting as the equivalent of a two-to-four wire hybrid transformer in a telephone. This duplexer is a passive three-port designed to pass energy in the receive frequency band from the antenna port to the receive port, while attenuating energy in the transmit band from the transmitter into the receiver. It is either made with high-dielectric ceramic resonators [51], or with a SAW filter and coaxial resonator in parallel [52]. The receive portion of a conventional handset resembles a broadcast

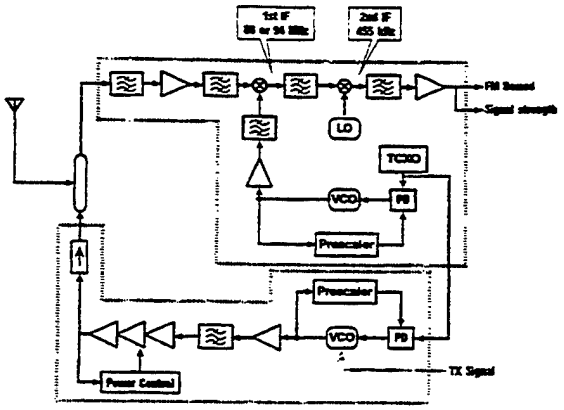


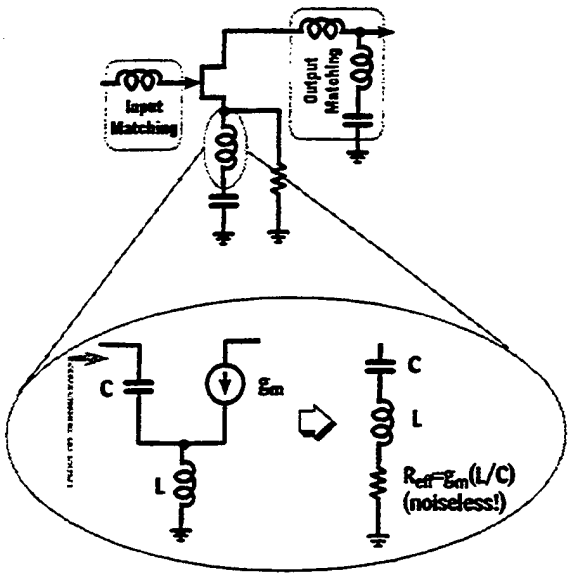
Fig. 9. The transmit and receive front-end of an analog cellular telephone. Receiver uses conventional double-superheterodyne architecture, while transmitter is one-step upconversion.

FM receiver (Fig. 9). If the local oscillator at the first mixer lies at a higher frequency than 894 MHz, the image is guaranteed to lie outside the AMPS band. Furthermore, an LO offset of more than 45 MHz ensures that the image is attenuated by the receive-band SAW filter, suffering at least 20 dB loss in each of the two filters in the handset [53]. A first IF of 90 MHz is therefore often used; this also avoids problems caused outside the handset by parasitic LO leakage through the antenna [54]. The desired channel is selected by locally synthesizing the first IF. This is followed by downconversion to a fixed second IF of 455 kHz, then demodulation by a frequency discriminator.

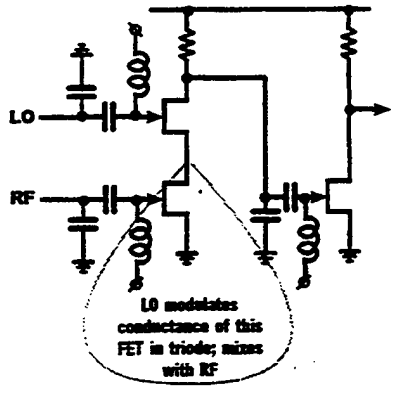
The first-generation of small-scale IC's for portable communication devices offered building-blocks for the intermediate-frequency chain, such as the mixer and local oscillator for conventional single or double-superheterodyne receivers [55], the IF amplifier chain and signal-strength indicator [55], or a standalone image-reject mixer [56]. Today, almost every major semiconductor company with an interest in the communications market offers building-block IC's at this scale of integration.

Although IC's entered the IF portions, the RF front-end circuits continued to be made from discrete components. An RF amplifier and a first mixer may be mounted with the associated filters on a dense miniature board [57], [58]. Discrete bipolar transistors responded to needs for portable RF applications, by offering, for instance, a low noise figure and f_T exceeding 5 GHz at less than 1 mA bias currents [59], [60]. Manufacturers of passive components, too, have steadily scaled-down their package sizes for high-density board mounting.

There is little argument, though, that the RF front-end components must also be integrated to reduce power dissipation. In competing with passive solutions, the RF-IC's must cross some important thresholds of low-price and high-performance [61]. However, they offer the prospect of an order-of-magnitude reduction in physical volume of the front-end electronics, and power savings will accrue by routing RF signals at a high impedance on-chip, while eliminating the low characteristic impedance interconnects



(a)



(b)

Fig. 10. Typical standalone GaAs MMIC's implementing (a) Low noise amplifier and (b) downconversion mixer. Concept of series-feedback inductor for low-noise matching illustrated.

between discrete packages. Various reasons have been advanced for why GaAs MMIC technology is now the right choice for cellular telephones [62]–[64]. They are summarized as follows: first, that owing to the semi-insulating nature of the GaAs substrate, reasonable-size inductors may be integrated with transistors to make high-frequency monolithic tuned circuits, which allows for lower current operation at a given frequency than would be possible with RC broadbanding techniques; and, second, that MESFET's afford lower noise figures at a given bias current than a bipolar transistor in a comparable silicon bipolar technology. Most GaAs MMIC's integrate front-end components for cellular applications at a small scale. For instance, a chip may integrate a tuned RF low-noise amplifier in the 900 MHz band, or a mixer and a local oscillator [64]–[72].

The typical RF amplifier may consist of only one or two FET's, with LC matching circuits on the input and the

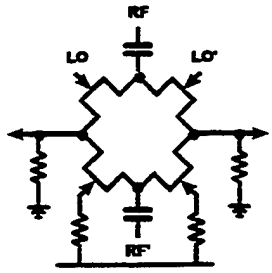


Fig. 11. A four-FET commutating switch mixer. The balanced LO signal switches the FET's, which chop the RF signal to produce the sum and difference frequencies.

output ports for standalone operation in a 50 Ω environment (Fig. 10(a)). A powerful and popular method to match the capacitive FET input is to insert a series feedback inductor, L, in the FET source, which at high frequencies contributes a *resistance* $g_m L/C_{GS}$ at the input port [73]. This method is preferred to resistive feedback found in wideband amplifiers for impedance matching, because unlike feedback resistors, the inductor does not degrade the noise figure. MMIC's from Matsushita favour the use of dual-gate MESFET's with RC matching instead of inductors [74], possibly because spiral inductors consume too large a chip area. GaAs IC designers must closely watch their chip real-estate to remain competitive in price. The various low-noise amplifiers operate in the 0.9–2 GHz bands, with gains of 15–20 dB, and noise figures of around 3 dB. These submicron MESFET IC's drain anywhere from 3–5 mA.

The received RF is typically downconverted by the local oscillator modulating the conductance of the mixer FET through a cascode FET (Fig. 10(b)). The mixers yield conversion gains greater than 10 dB, noise figures of 10–12 dB, and third-order input-referred intercept points of −5 to 0 dBm. Dual-gate MESFET's are naturally suited for mixer use, and offer a similar performance [74]. A recent 0.7 μm GaAs MMIC offers an LNA-mixer pair draining 3 mA from 3 V, with the LNA producing 13 dB gain, 3.6 dB noise figure, and an input-referred IP3 of −11 dBm [75].

Slicing across the system a little differently, another MMIC implements the downconversion and upconversion mixers for the receiver and transmitter, respectively, integrating their shared local oscillator on the same substrate [76]. The four-FET switch mixer (Fig. 11) is very linear, but it suffers from two disadvantages: it requires a large local oscillator drive to turn the switches on and off [77], and unlike the bipolar transistor Gilbert-cell analog multiplier, the switch mixer is lossy, requiring additional signal amplification from the following stages.

The current generation of analog cellular telephones transmits power levels of more than 1W (30 dBm). RF power amplifiers are usually packaged in a separate module with some form of integral heat-sink. A preamplifier, or in radio terminology, exciter, in the transmitter section boosts the modulated carrier level close to 0 dBm to drive the power amplifier input. The power amplifier module itself usually consists of a cascade of two or three FET's, tapering up to a single large-size FET which will deliver the

required signal current into the antenna load. Furthermore, as the power amplifier is the largest single source of battery drain, it must have a *high conversion-efficiency*. Much as in baseband power amplifiers, an efficient RF power amplifier is biased close to cutoff to reduce the dc standing current, and then driven by the input signal in Class A-B or Class-B mode. Narrowband filters at the amplifier output remove harmonic distortion caused by nonlinear operation at RF [78]. These filters may be merged into the passive matching networks required for optimum power transfer from the amplifier to the load. As designers of RF power amplifiers have observed [10], [79], [80], much of their work consists of synthesis and iteration of the interstage, input, and output *matching networks*. It is easier to design efficient power amplifiers for constant-envelope modulations, such as analog FM or digital FSK, where distortion may be tolerated because the useful information is all contained in the zero crossings. Further, it is argued that owing to the lower parasitic capacitance of a GaAs MESFET relative to silicon devices, a GaAs power amplifier at 1W power levels affords a higher efficiency (~60%) compared to silicon bipolars or FET's (~ 45%) [47]. A monolithic power amplifier consisting of a four-stage cascade of MESFET's with on-chip lumped LCR input and inter-stage matching networks, delivers 1W at 900 MHz at 63% efficiency from a 5.5 V supply [80]. The distributed element output matching network at 900 MHz would be exhorbitantly large on an IC, and is therefore printed on an off-chip alumina substrate. The high efficiency is attributed to an improved method of suppressing the 2nd harmonic at the amplifier output. This multi-component module approach to mate IC's with matching networks is widely used in GaAs power amplifiers [81]. Another 900 MHz power module uses two discrete MESFET's, wirebonded to a hybrid IC containing a combination of distributed circuits and chip capacitors and resistors in the matching network, to attain 65% efficiency when delivering over 1.3 W from a 4.7 V supply [82]. The output FET is 12 mm wide, with a 1 μm channel length. A separate negative supply is required in many of these MESFET power amplifiers to bias the gate. The circuit was recently modified [83] to produce the same output power equally efficiently, now with a single 3.5 V supply and FET's of 0.6-μm channel length. The desired low-voltage operation, much more suitable for one battery, is the result of an improved device structure. The matching networks in the RF signal path are fabricated on separate passive-only GaAs IC's, which the authors claim are three times cheaper than GaAs IC's with FET's, while the bias networks reside on a miniature PC board. The various components are wirebonded to one another, and mounted on an AlN substrate with ten times higher thermal conductivity than alumina. Power amplifiers may also be built with silicon NMOSFET's: among the major semiconductor vendors, Hitachi has pursued this option. A MOSFET gracefully accepts large voltage swings on the gate, without possibility of Schottky conduction as in a MESFET gate. With an offset-gate FET structure and 0.8-μm channel length, a MOSFET power amp delivers 2 W

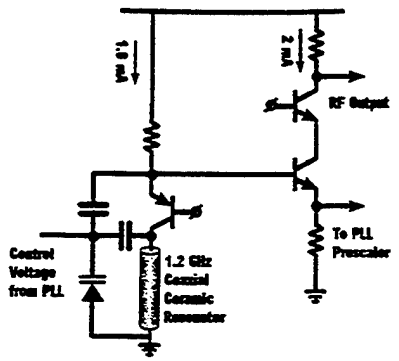


Fig. 12. Colpitts oscillator with coaxial resonator. Varactor controls frequency with control voltage. Series-connected buffer distributes RF oscillation to mixer and to PLL prescaler.

at 1.5 GHz with 55% efficiency from a single 6-V supply [84]. The power amplifier module embeds these FET's into matching networks [85].

The foregoing discussion is not meant to imply that cellular telephones mainly use GaAs MMIC's in the RF front-end. Silicon bipolar and BiCMOS technologies have made tremendous strides in improving f_T, and today silicon MMIC's offer comparable performance to what is available in GaAs. A recent 1 GHz BiCMOS LNA-mixer combination [86] uses familiar circuits to baseband designers, such as a Gilbert-multiplier type mixer, and the chip yields a comparable performance to the MMIC's described above. The RF signal path consists of bipolar circuits only, while the FET's are used as switches to select various power-down modes. There are no on-chip inductors to tune the low-noise amplifier, which is wideband; instead, the LNA output is routed off-chip into a passive bandpass filter, then returned to the chip for downconversion. This is part of a complete chipset from Philips for a digital cellular telephone handset [87].

The first LO in a cellular telephone receiver is programmed to the incoming RF channel with a phase-locked loop synthesizer, while the second LO at IF is fixed in frequency. The transmit LO oscillates in yet another frequency band. Both the receiver and the transmitter therefore require 900 MHz voltage-controlled oscillators. These are most often implemented with bipolar transistors, whose very small flicker noise means low phase-noise sidebands in the VCO. A stripline resonator sets the nominal frequency of a Colpitts oscillator, and this is voltage-controlled by a varactor diode in parallel [54], [57], [88]. Phase noise levels of -110 to -120 dBc/Hz are attained at a 50 kHz offset from the oscillation frequency. The VCO application has created a brisk demand for varactors with large voltage-coefficient and low-loss [60]. The VCO and its buffer may be connected in series to reuse the bias current between the two stages [57, [88] (Fig. 12). The nominal oscillation frequency may be slaved in a frequency multiplying PLL to a 12 or 15 MHz crystal oscillator.

The Motorola MicroTac, first introduced in 1989, set an industry standard for a miniature cellular hand set. Small

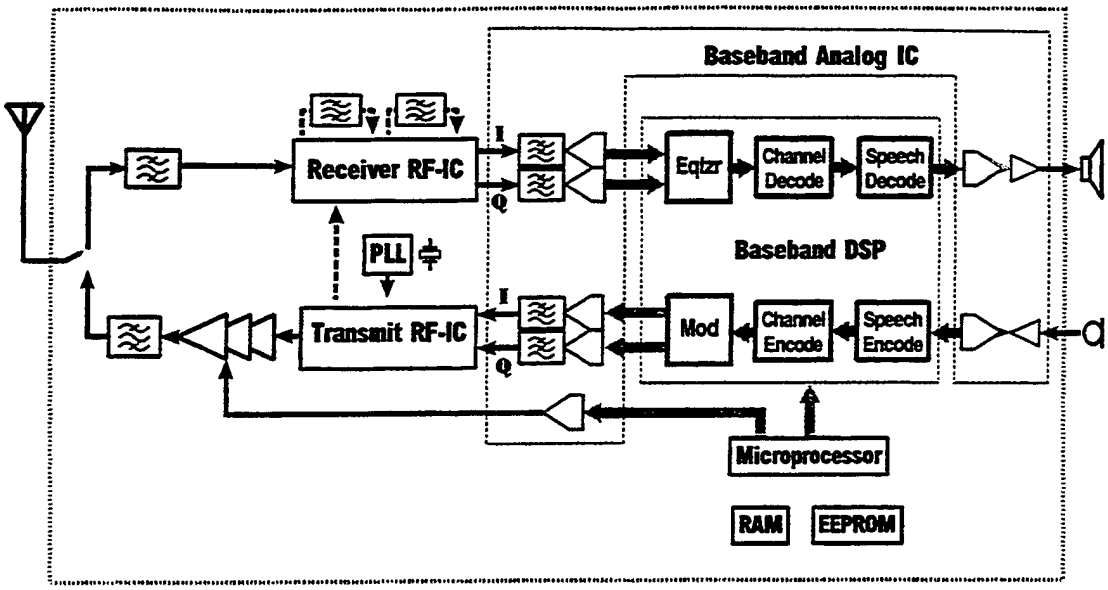


Fig. 13. Block diagram of digital cellular telephone. Chip partitions reflects current state-of-the-art technology.

handsets [89] continue to use conventional architecture, but attain a small size with discrete filters in miniature packages [52], [53], [58], [90], [91], and with lower power electronics which reduces battery weight. The electronics entail carefully designed standby modes, better software for the on-board microprocessors, and, among other items, lower power PLL-prescalers, digital counters which must operate all the time for frequency synthesis and which often posed a significant power drain [92], [93]. In the transmit mode, greater efficiency is sought in the power amplifier to prolong battery life, as well as smart control circuits to maintain maximum efficiency across the range of output power levels [89]. It was generally agreed that by 1992 the handset volume should scale down to less than 150 cc, and that its weight should be less than 230 gm [20], [94]. NTT demonstrated the prototype of such a telephone in 1991 [95], while noting that the volume should be no smaller than this target both for the sake of ergonomics, and to ensure adequate heat removal from the package so that the internal temperature rises no more than 15°C. These small telephones might be powered by a 6-V NiCd battery with 400 mA · h capacity.

With the emergence of the IS-54 standard, there is some convergence in the US on a digital cellular technology. Upward compatibility is sought with the existing analog cellular bands, which is why handsets conforming to IS-54 are referred to as dual-mode cellular. A recent 12 GHz f_T silicon bipolar transceiver IC from AT&T uses a conventional double-superheterodyne architecture with an 80 MHz first-IF, and after IF amplification the signal path bifurcates into a conventional 455 kHz second-IF FM demodulator, or a digital *I-Q* PSK detector [96]. An interesting alternative is to downconvert the second-IF, without preceding AGC, in a delta-sigma A/D converter for subsequent digital baseband signal processing [97].

VI. Digital Cellular and Cordless Telephones

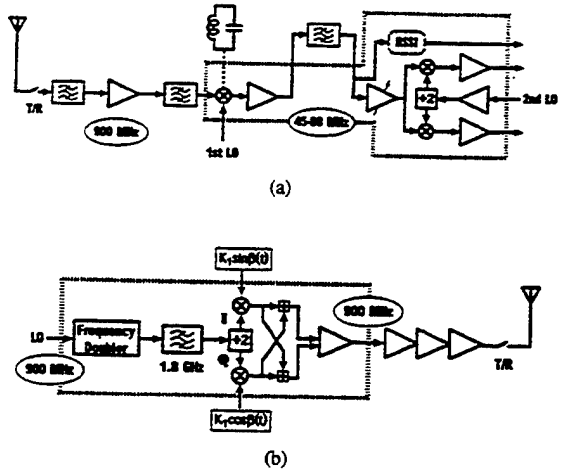
Analog cellular telephones use the frequency spectrum inefficiently. The modulation schemes consume a large bandwidth, and every cellular telephone transmits at constant power all the time it is in use, thereby appearing as an interferer to other users at nearby frequencies. As a result, the spectrum allotted to cellular phones in large metropolitan areas nears exhaustion.

Digital cellular telephones are one solution to better utilize the scarce spectrum. They use more efficient modulation schemes, such as minimum frequency-shift keying or phase-shift keying, and multiple users may share time-slots on the same part of the spectrum. The complex modulation formats used in these telephones and the greater capabilities required to withstand nearby blocking signals are prompting large-scale integration of the RF and IF electronics. Examples of digital telephony standards are the European GSM, North American Digital Cellular, and the emerging Japanese personal handy phone (PHP) [98]. The salient charecteritics of these various systems, as well as key digital cordless standards, DECT and CT-2, are summarized in Table 1.

The typical handset involves an RF/IF front-end, followed by a baseband digital signal processor (Fig. 13). The first significant set of RF-IC's for GSM handsets appeared in 1990. Notable among them is a receiver and transmitter 7.5 GHz f_T silicon bipolar chip-set from Siemens (Fig. 14) [99], [100]. The double-superheterodyne receiver uses a selectable first-IF of anywhere from 45–90 MHz, which places the image channel in the stopband of the 25 MHz-wide RF preselect bandpass filter. A 71 MHz IF, for instance, guarantees that the image lies outside the GSM band. The desired 200 kHz channel is selected by a SAW filter at the first IF. The receiver IC provides an on-chip

Table 1

	GSM Europe	NADC North America	J-PHP Japan	CT-2 Europe, Asia	DECT Europe
Downlink Frequency Band	935–960 MHz	869–894 MHz	1.9 GHz	864–868 MHz	1.88–1.9 GHz
Uplink Frequency Band	890–915 MHz	824–849 MHz	1.9 GHz	864–868 MHz	1.88–1.9 GHz
Multiple Access Method	TDMA	TDMA	TDMA/TDD	FDMA/TDD	TDMA/TDD
Modulation	GMSK	$\pi/4$-DQPSK	$\pi/4$-QPSK	B-FSK	GMSK
Speech data rate	13 kb/s	8 kb/s	32 kb/s	32 kb/s	32 kb/s
Handset output power	3.7 mW → 1W	2.2 mW → 6 W	10 mW	1 mW → 10 mW	250 mW
Modulation rate	271 kb/s	49 kb/s	384 kb/s	72 kb/s	32 kb/s
Channel spacing	200 kHz	30 kHz	300 kHz	100 kHz	1.762 MHz
Burst length	156 b	324 b			424 b



(a) (b)

Fig. 14. Siemens receiver and transmitter IC's for GSM handsets. (a) Receiver is double-superheterodyne architecture. Requires one off-chip tuned circuit, SAW resonator at first IF, and low-noise amplifier. (b) Transmitter doubles LO frequency and then divides-by-2 to accurately obtain quadrature phases.

buffer to drive this filter. Following an RF amplifier with an off-chip LC tuned load, the input signal is downconverted by a Gilbert-cell mixer. There is AGC with 70 dB range and a signal-strength indicator at the first IF, and the selected channel is then downconverted to baseband in a quadrature demodulator to acquire both amplitude and phase for GMSK vector demodulation. The on-chip gain may be as large as 80 dB at the highest VGA setting, but as it is distributed in different frequency bands, there is little on-chip crosstalk and the monolithic receiver operates stably. The receiver uses a fully balanced signal path, and an input noise figure of 7 dB. It drains 27 mA from 5 V in active mode, and 10 μA in standby mode.

In the transmitter IC, a precision quadrature upconversion mixer produces a single-sideband, suppressed-carrier QPSK output from a baseband vector input. An RF differential input amplifier selects one sideband from the upconverted outputs in the two arms of the mixer. As the other sideband would occupy a nearby channel's spectrum, it must be adequately suppressed (relative to the level of the wanted sideband). Imperfect cancellation of the unwanted sideband arises from gain mismatch in the I and Q channels of the upconversion mixer, and in deviations from quadrature in the two LO outputs. The goal in this transmitter was quadrature phase-errors of less than 2°, which with adequate gain matching, implies unwanted sideband suppression of at least 40 dB relative to the wanted sideband. The stringent GSM requirements on low LO phase noise are only met with a 900 MHz off-chip coaxial-resonator based, varactor-tuned, Colpitts oscillator [101], which inherently provides a single-phase output. There are several methods to accurately derive quadrature phases. Siemens uses the double-frequency method. An on-chip nonlinear element doubles the oscillator frequency to 1.8 GHz, following which an on-chip bandpass filter removes harmonics. Then, the oscillation is divided by 2 by positive- and negative-edge triggered flip-flops, which will yield quadrature outputs with a phase-accuracy set by how close the duty-cycle of double frequency oscillation is to 50%. Combined with an output stage to drive an off-chip power amplifier module, the IC drains 40 mA. Using a similar superheterodyne architecture with 71 MHz IF, AT&T Microelectronics has recently combined the GSM receiver and transmitter blocks, including the frequency synthesizer, on to one chip [102].

In an effort towards even greater miniaturization at Alcatel, the transmitter and receiver sections are both integrated on to one 9 GHz f_T silicon bipolar chip (Fig. 15) [103], [104]. From the point of view of integration, the main advantage of the zero-IF, direct-conversion architecture is that it eliminates the IF passive high-frequency bandpass filters. The disadvantage, on the other hand, is that the downconverted signal has energy at dc, to which will add the receiver's offsets and low-frequency noise. An off-chip low-noise amplifier drives directly into quadrature (I-Q)mixers. The mixer dynamic range is wide enough so that a large blocking signal only 3 MHz away from the desired signal produces insignificant intermodulation distortion. The bipolar Gilbert-cell mixer is linearized with emitter degeneration resistors, which also degrade its noise figure. The receiver section requires large (650 pF) off-chip capacitors at its output for anti-alias lowpass filters, before the baseband signal is sent to a companion mixed-signal CMOS IC. This baseband CMOS IC contains a high-order switched-capacitor lowpass filter for channel selection, a digital GMSK detector, and various DSP functions [105]. An RC and CR network with off-chip trimming shifts the phase of an external local oscillator by precisely ±45° to produce the quadrature drive to the mixers. Upconversion mixers in the transmitter drive a power amplifier module.

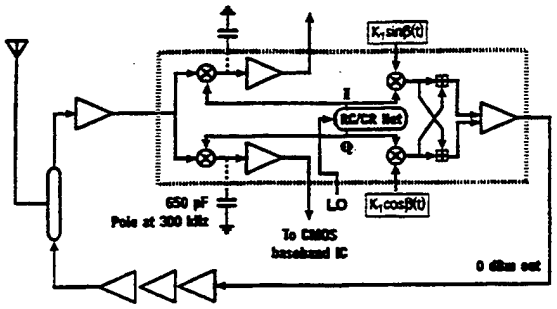


Fig. 15. Alcatel direct-conversion single-chip transceiver. Channel selection done by switched-capactior lowpass filters in companion CMOS mixed-signal baseband IC. On-chip, trimmed RC-CR network generates quadrature phases for mixers.

The chip drains 25 mA from 5 V in the receive mode, and 45 mA in the transmit mode, rather comparable figures to the previously decribed Siemens chip set. The transmitter suppresses the unwanted sideband and carrier by about 40 dB.

There is increasing interest in the 1.9 GHz band for digital cellular telephony. All the current approaches (Table 1) use time-divison duplexed (TDD) receive and transmit frames, wherein users are assigned different time slots. Two-way communications take place over the same frequency band. A host of IC's is appearing to serve the European DECT standard. The majority cater to a superheterodyne receive-architecture with 110 MHz IF [106]–[109]; some of the literature [107], [109] also describes procedures for system design. The RF sections are usually integrated at small scales, embodying, say, the low-noise amplifier and first mixer on one chip, the quadrature modulator on another, and the exciter and power amplifier on a third and fourth chip. RF-IC's operating at 1.9–2.5 GHz exist in both GaAs [68], [110]–[112] and silicon [106], [113]–[115] technologies. Some of these standalone GaAs IC low-noise amplifiers achieve impressive gain and noise figure with 1–2 mA current drain from 3 V [69], [116]. At the system level, though, functionality and overall dynamic range in these short-haul wireless links takes precedence over raw component performance. Thus system-level input noise figures of 10–15 dB are acceptable [109], as are input-referred intercept points of −16 dBm, and a 20 dB power amplifier control range [112]. The less than stringent system specifications lead to simplifications in the transmit path, such as direct VCO-modulation by the baseband signal by opening the transmit PLL over the duration of the transmit-frame [107].

The power amplifier continues to be built almost exclusively as a separate GaAs IC, and at 1.9 GHz the on-chip wavelength is short enough that it is now possible to integrate distributed matching networks [68], [117], [118]. Efficiencies of 50% are attained when delivering almost 1 W to the load from 3 V. Either the transmitter or the receiver is active in TDD digital transceivers, so fast-switching power-down modes are designed into the various components. In particular, the switching trajectory must be shaped to suppress spurious emissions when the power amplifier is switched on and off [85]. Finally, the transmit and receive signals are directed to the antenna via a three-port passive circulator, or preferably through a low-loss, monolithic transmit/receive (T/R) RF switch [112], [119], [120]. A good FET switch must not appreciably distort the RF signal, or incur more than 1-dB insertion loss when ON, yet when OFF it should offer at least 30 dB isolation. Microcell applications, such as the Japanese handy phone, require an average output power of only 10 mW (20 dBm), although the $\pi/4$-QPSK modulation requires the power amplifier to handle larger peak powers. An exciter amplifier, power amplifier, and T/R switch have been integrated together [121]. The exciter and power amplifier attain a 44% efficiency at 23 dBm output from a 4.8-V supply, and the on-chip matching network uses LC lumped elements. Although matching networks will eliminate harmonic distortion, they are ineffective in suppressing near-carrier spurii produced by intermodulation distortion. Feedback linearization techniques have been proposed which predistort the digitally synthesized exciter input waveform to anticipate the power-amplifier nonlinearities [122]. Use of these techniques makes a highly nonlinear but efficient power amplifier appear linear. These remain at the experimental stage today, and with the move to microcells and low emitted power levels, they may soon not be necessary in handsets.

The building-block IC's described so far are important advances in realizing small, relatively low-power transceivers, but the ultimate goal remains to integrate the entire transceiver on to a single-chip. To this end, some early breadboard-level experiments show that direct-conversion receiver architectures seem well suited to the DECT application [123], [124]. A 16 GHz f_T silicon bipolar IC from Alcatel operating at 1.9 GHz contains a complete direct-conversion DECT transceiver [125]. The architecture resembles the previously described Alcatel direct-conversion GSM transceiver [103], requiring a separate low-noise amplifier in the receive path, a power amplifier module, and a baseband CMOS mixed-signal signal processor IC. From a 5-V supply, the transceiver drains 50 mA in receive mode and 80 mA in transmit mode. One of the challenges in realizing this system was a fast-switching PLL frequency synthesizer with low phase-noise and low spurious output levels, which was built here with an improved charge-pump and loop filter.

Siemens has extended its dual-conversion GSM transmitter/receiver chip set [99], [126] to a generalized front-end for digital cellular telephones operating in any part of the RF spectrum from 800 MHz to 2.1 GHz [127]. The 25 GHz f_T silicon bipolar chip set ope rates at a supply as low as 2.7-V, the transmitter chip draining 60 mA and the receiver 33 mA. Both have power-down modes. An external power amplifier module is required. A notable low-voltage circuit on this chip is a modified Gilbert-cell mixer, with resistors instead of current sources to the negative supply [31]. To accommodate a wide input dynamic range, the gain of the RF low-noise amplifier in the receiver may be switched from −5 dBm to +15 dBm, a powerful technique

that has also been used in paging receivers [46]. This assumes that the recever does not instantaneously require a very wide dynamic range, but either receives mostly strong signals or mostly weak ones. The downconversion mixer, with a 13 dB noise figure and −3 dBm input IP3, is followed by an IF variable-gain amplifier with an 80 dB digitally programmable range [126]. The on-chip RF oscillator requires an off-chip resonator and varactor, and its phase noise at a 2 kHz offset is −88 dBc/Hz. A balanced signal path is used throughout. One good measure of the quality of on-chip gain matching in the I and Q paths in the direct upconverter, as well as of the quadrature accuracy of the upconversion clocks, is given by the relative levels of spurious emissions from the transmitter IC. The unwanted sideband is suppressed by 48 dB, while the RF carrier tone, emitted due to dc offsets in the two paths, is 37 dB down. A 3rd-order modulation tone at −46 dB appears due to mixer nonlinearity. The absolute output noise level at a 25 MHz offset is −141 dBm/Hz.

VII. Direct-Conversion Transceivers and Their Problems

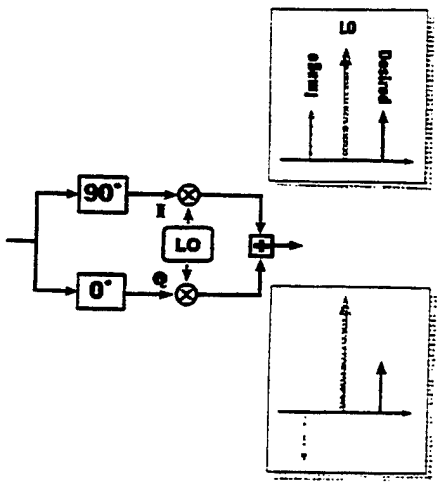
A receiver with zero-IF is called a *direct-conversion* receiver. When the local oscillator is synchronized in phase with the incoming carrier frequency, this is also referred to as the *homodyne*. This architecture is sufficiently promising for single-chip transceivers to warrant a separate section to its study.

The desired channel is translated by the first mixer to a 0 Hz center frequency, and instead of adjacent channel rejection with a bandpass resonator, a more flexible and easier to implement lowpass filter is required—in effect, a bandpass filter centered at dc when the negative frequency axis is included. With zero-IF, there is *no* image frequency. As early as 1924, radio pioneers had considered homodyne architectures for crude receivers requiring only a single vacuum-tube, but it was in 1947 that a homodyne was first used to full effect, with a high-order lowpass filter for channel-selection, in a measuring instrument for carrier-based telephony [32].

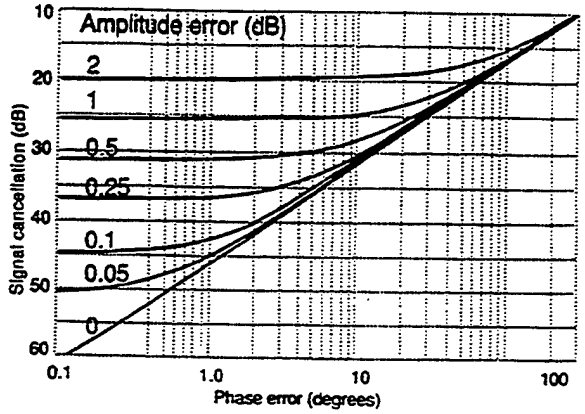
A. Direct-Conversion Single-Sideband Synthesizers

For reasons of spectral efficiency, the transmitted signal in digital communications is always single-sideband with suppressed carrier. This is most often produced with the so-called *phasing method* [128]. The modulated signal is first synthesized in quadrature at baseband, *directly upconverted* into two paths by a quadrature LO centered at the carrier frequency, and added or subtracted to select either the upper or lower sideband (Fig. 16(a)).

The unwanted side band is suppressed to an extent limited by the gain mismatch in the two upconversion paths, and by departures from quadrature in the two LO outputs (Fig. 16(b)). Unequal dc offsets in the paths produce an output signal at the LO frequency. The unwanted sideband and LO leakage are spurious but unavoidable components of the transmitted spectrum. Although on the same IC the



(a)



(b)

Fig. 16. (a) Single-sideband direct upconverter using the phasing method (also called the quadrature modulator). Extent of suppression of unwanted (image) sideband versus gain mismatch and phase-errors from quadrature in two arms of modulator.

two upconversion paths are well matched, a gain mismatch as small as 1% (0.1 dB) will limit suppression of the unwanted sideband to 45 dB. With this gain mismatch, a phase-error of up to 1° may be tolerated between the two LO outputs. These mismatches may be trimmed down at time of transceiver manufacture, or self-calibrated with loopback modes which are activated during idle times to sense and suppress the unwanted spurii.

An off-chip resonator, often connected to an on-chip unbalanced oscillator circuit, may also become a cause of spurious RF leakage if it couples energy into the power amplifier or the antenna. Frequency-offset upconversion schemes have been proposed [129] to combat this coupling problem. Other spurious output tones may arise from parasitic remixing of the modulated output with the baseband signal, and by intermodulation distortion in the output stage [96]. Balanced circuit topologies, on-chip LO's requiring no external resonators [130], and lowered transmit power levels in microcells, are all expected to lessen the magnitude of the spurious leakage problem.

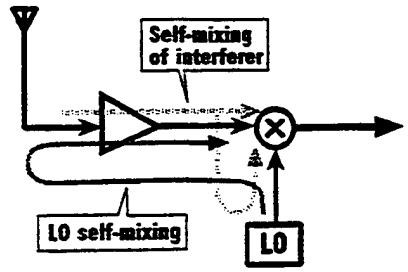


Fig. 17. Sources of leakage in a direct-conversion receiver, and how through self-mixing they create dc offsets.

Cellular wireless systems operate best with tightly regulated power control, the handsets transmitting only the minimum power required for reliable reception by the base station, and *vice-versa*. Whereas in the receiver the front-end is the main source of noise, in a transmitter the upconversion LO phase-noise appears as added noise on the emitted signal. Direct upconversion has the advantage over two-step schemes that only *one* LO contributes noise.

B. Direct-Conversion Receivers for GSM Digital Cellular Telephones

Among the various RF-IC suppliers for the European GSM digital cellular telephone handset, only Alcatel at present uses a direct-conversion receiver architecture [103], [125], [131]. This results in a relatively small silicon bipolar RF front-end chip, and the remainder of the signal processing, including lowpass channel-select filtering, is at baseband in mixed-mode CMOS. Why are the others reluctant to use direct-conversion in their receivers? Given the many decades of superheterodyne experience, the most likely reason is conservatism. But direct-conversion also suffers from some unique problems.

A well known problem is that spurious LO leakage from the receiver into the antenna becomes an in-band interferer to other nearby receivers tuned to the same band. Superheterodynes, with their frequency-offset LOs, do not suffer from this problem. However, experimental studies suggest that with standard shielding in the receiver, this problem is not so severe as to handicap the use of direct-conversion [132].

A more serious problem is dc *offset* in the receiver. Offset arises from three sources [131]: transistor mismatch in the signal path; the LO signal leaking to the antenna because of poor reverse isolation through the mixer and RF amplifier, then reflecting off the antenna and self-downconverting to dc through the mixer (Fig. 17); and a large near-channel interferer leaking into the LO port of the mixer, then self-downconverting to dc. Good circuit design may reduce these effects to a certain extent, but they cannot be eliminated.

The spectrum of the GMSK modulation used in GSM has a peak at dc. Offsets will directly add to the spectral peak of the downconverted signal. These offsets are usually much larger than the rms front-end noise, and may therefore significantly degrade the SNR at the detector. To remove the offsets by ac-coupling the receiver will require impractically large capacitors, if the signal-bearing spectrum around dc is not to be sacrificed. However, they may be compensated with DSP-based self-calibration [105], [133]. The baseband signal processor makes long-term measurements on the dc level in the receiver, and subtracts off this level from the downconverted signal. The spectrum loss around dc is only a few hertz, and digital filtering does not distort the midchannel group delay in the receiver.

C. Local Oscillators with Quadrature Outputs

A *carrier-frequency* local oscillator with quadrature outputs is a key circuit component in direct-conversion transmitters and receivers. Usually this oscillator is tuned by an LC circuit or an off-chip resonator, and inherently produces a single-phase output. Quadrature phases are often derived by passing the oscillator output through a CR and an RC network, whose time constant is equal to the oscillation period. The two resultant outputs are then phase-shifted by $+45°$ and $-45°$, respectively. Inaccuracies in the actual values of R and C will lead to errors in quadrature, and are compensated by some form of on-chip trimming [103], [125].

An alternative method is to divide an oscillation at twice the carrier frequency with positive- and negative-edge triggered flip-flops [99]. The resulting two outputs at the desired frequency are in quadrature. Residual phase errors caused by unequal delays in the two flip-flops, and by departures from 50% duty cycle in the double-frequency oscillation, are found to be less than 1°. However, the double-frequency portions of the circuit may become a speed bottleneck.

Quadrature outputs may also be derived from a polyphase oscillator, such as a variable-frequency ring oscillator. In a four-delay stage ring oscillator, taps at diametrically opposite points will yield quadrature phases at all frequencies [134] (Fig. 18). In oscillators with an odd-number of unit delays, they may be synthesized from two taps by a voltage-controlled phase-shifter in feedback around a quadrature-sensing circuit [135]. Mismatches in transistor characteristics limit the attainable phase accuracy. This type of resonator-less oscillator is practical when the specifications on phase-noise are not too stringent (say up to -80 dBc/Hz at a 100 kHz offset from the carrier). The free-running oscillator must be slaved to a crystal reference in a frequency-locked loop, whose loop gain and bandwidth are specifically designed to suppress phase noise.

When circuit techniques cannot contribute further improvements to gain- and phase-matching in the two arms of a quadrature modulator or demodulator, digital adaptive algorithms may be used at baseband to sense and compensate for these errors [136].

VIII. IC's for Spread-Spectrum Wireless Transceivers

Spread-spectrum communications offer greater user capacity than narrowband techniques in a given piece of

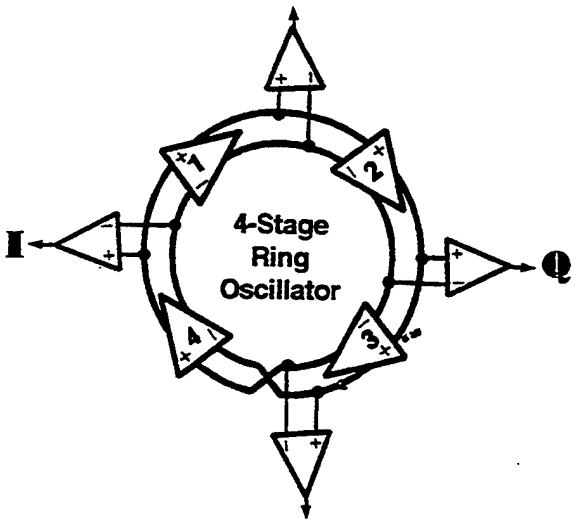


Fig. 18. A four-phase voltage-controlled ring oscillator. Diametrically opposite taps give quadrature phase at any frequency. Unequal loadings and device mismatch set phase errors. VCO must be embedded in PLL whose closed-loop bandwidth is tailored to suppress oscillator noise.

wireless spectrum. Much has been written on this subject [137]. Spread-spectrum techniques were developed during World War II as a form of secure communication with low probability-of-intercept and resilience to jamming [138]. Over the years, this technology has been further developed and refined for military communications. With the wireless revolution at hand and the IC technology now available to implement complex transceivers, spread-spectrum has awakened commercial interest [139]. Spread-spectrum communications make it easy for users to access the wireless channel. Whereas conventional narrowband wireless communications require a careful discipline, which the FCC or other government agencies enforce by issuing licenses so as to prevent use of the same frequency by nearby operators, for spread-spectrum use the FCC has allocated certain *unlicensed* bands in the US, referred to as the Instrumentation, Scientific, and Medical (ISM) bands, wherein the user is only required to spread spectrum by a minimum amount, and not to exceed an upper limit on transmitted power. The extent of spectrum spreading may be measured by the amount of *processing gain* in the receiver required to de-spread and detect the signal [140].

As analog cellular telephones begin to saturate the available radio frequency allocations, spread-spectrum techniques are being deployed in cordless telephones and wireless modems [141]. These devices usually operate in the two lower ISM bands: 902–928 MHz and 2.4–2.48 GHz. A ubiquitous wireless environment is envisioned, in which mobile users, wherever they may be, can access data and communications services through an intricate network of base stations [2]. The greatest hardware challenge in realizing this scenario is the development of a low-power, miniature handset.

There are two different methods to spread the spectrum of a signal: by *direct-sequence* modulation, or by *frequency hopping* [140]. Direct sequence is conceptually the simpler, as well as the more straightforward to implement. Each data bit (either $+1$ or -1) at the transmitter multiplies a prescribed sequence of bits, or *chips* as they are called. The chip sequence is selected for very low autocorrelation, and is often referred to as a pseudo-noise, or PN, sequence. Each user is either assigned a unique PN sequence with very low cross correlation with other user sequences, or a time-shifted version of some long PN sequence. The receiver, after an initial acquisition search to align itself with the start of the PN sequence, correlates the incoming sequence with the pattern it knows to be its own. On detection of a correlation peak, the sign of the peak signals the source data bit. The longer the PN sequence for each bit, the greater the spreading factor, and the more reliable the detection. Here, the processing gain is the sequence length per bit. The spectrum of the transmitted data, composed from a concatenation of PN sequences, is noise-like; thus the term *spread-spectrum*. Several baseband and IF modem IC's have already appeared to support BPSK and QPSK direct-sequence spread-spectrum transceivers [142]–[144].

Whereas direct-sequence modulation spreads the spectrum by randomizing the waveform, frequency-hopping spreads spectrum by *hopping* the carrier according to a

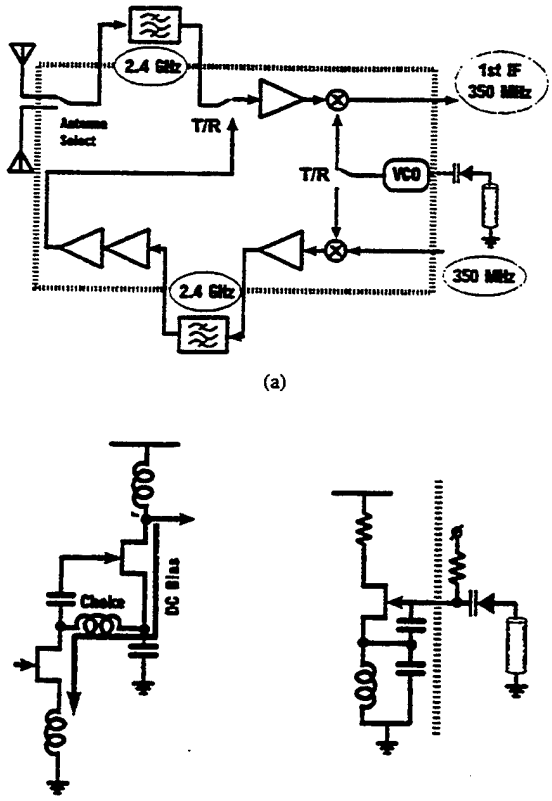


Fig. 19. Front-end GaAs MMIC transceiver for spread spectrum communications in 2.4 GHz band. Receiver is double-superhetorodyne. Note use of off-chip filters. (b) Low noise amplifier in transceiver IC (left) reuses bias current in two stages. Clapp voltage-controlled local oscillator (right) combines on-chip LC network with off-chip resonator and varactor.

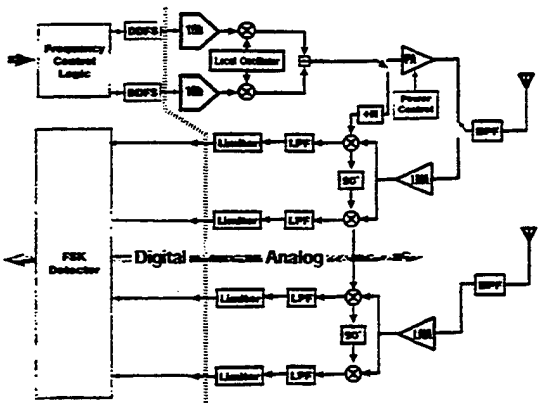


Fig. 20. The UCLA frequency-hopped spread-spectrum transceiver. This uses direct-conversion in both transmit and receive path, and binary-FSK modulation. In the transmitter, the spread-spectrum waveform is synthesized at baseband, and upconverted to RF. During receive mode, the DDFS under control of a synchronization loop hops with the incoming data, and demodulates it to baseband. The entire mixed analog-digital, baseband/RF system is to be integrated on a 1-μm CMOS IC.

prescribed sequence across the entire band. The effect in the frequency-domain is similar to direct-sequence, except for one fundamental difference: with direct-sequence, spreading across a wide bandwidth requires each data bit to be mapped into a long PN sequence. The resultant high "chip-rate" requires high-speed signal processing at the receiver front end. On the other hand, if the carrier frequency is hopped with a wideband synthesizer, the output spectrum may span an arbitrarily wide frequency range, even at low bit rates. Therefore, if a fast and agile frequency synthesizer is readily available, the receiver front-end operates at the actual data rate in frequency-hopped spread-spectrum, rather than at the considerably higher chip-rate in a direct-sequence receiver. The former will very likely lead to a low-power solution.

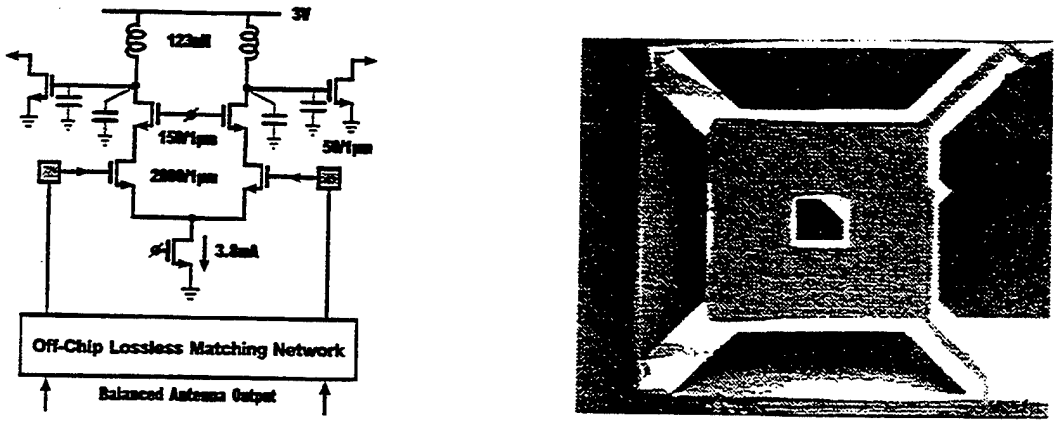
As a notable recent example of a miniature spread-spectrum transceiver, Plessey has introduced a 700 kb/s frequency-hopped device operating in the 2.4 GHz ISM band. The transceiver is entirely contained on a 2″× 3″ PCMCIA card for insertion into notebook computers [109], [145]. FSK data modulates the carrier frequency, and a variable-modulus PLL synthesizer slowly hops the carrier to spread the spectrum across the 80 MHz band. All the active devices in the transceiver are on three IC's, consisting of a GaAs RF front-end, a silicon bipolar IF receiver, and a CMOS IC for the hopping-frequency synthesis. As in any spread-spectrum two-way communication system, transmission and reception is time-division-duplexed on the same frequency band. The receiver architecture is a conventional double-superheterodyne. In addition, the transceiver requires 50 passive components, including six rather bulky filters, and when transmitting 100 mW RF power it dissipates more than 1 W.

By the norms of GaAs MMIC's, the single-chip IC front-end in this transceiver is highly integrated (Fig. 19(a)) [146]. It includes the power amplifier, and drivers for 2-GHz passive filters in the transmit and receive paths. Interesting features are the dc series connection of the single-ended two-stage low-noise amplifier, which shares the same bias current in both stages through a bypass inductor; the Clapp VCO; and the four-FET switched mixers in the transmit and receive paths (Fig. 19(b)). In its first version, the IC includes more than 20 on-chip spiral inductors, perhaps the largest number on any MMIC at the time of this design. A second bipolar IC processes the 350 MHz IF signal. In receive mode, the GaAs front-end IC drains 30 mA from a 5 V supply. The receiver selects one of two external antennas through an on-chip RF switch to attain spatial diversity. This work has spurred a flurry of similar GaAs IC's at the same level of integration [147]–[149], all operating at ±5 V supplies, with similar functionality and performance. One of them [149] uses a very high IF of 915 MHz, so that spurious products at the transmitter output lie well in the stopbands of the output filter, and where the IF signal is processed by a 915 MHz cellular-telephone type IC.

IX. IC's to Enable Future Transceivers

What will portable wireless communicators of the future look like? We may draw some conclusions from the foregoing summary of RF-IC developments. The crowded spectrum means that future wireless communicators will predominantly use spread-spectrum techniques. Coupled to this is a need for low-power dissipation, which will force *architectural innovations* and *higher levels of integration* in the electronics. As an example of such an advanced transceiver, the author with his colleagues and their graduate students at UCLA is investigating the architecture and circuit design of a frequency-hopped, binary frequency-shift keyed, zero-IF, all-CMOS two-chip transceiver capable of delivering up to 160 kb/s (the base ISDN rate) in the 900 MHz ISM band [39], [150]. The transceiver architecture is inspired by the modern paging receiver, a very low energy wireless device widely used today. The transceiver implementation (Fig. 20) will freely mix analog and digital circuits, which makes CMOS the IC technology of choice. Further, to avoid the routing of high-frequency signals off- and on-chip and thereby save the power the buffers would use to drive stray capacitance and off-chip low-impedance lines, it is preferred to integrate all the blocks, *including* the RF front-end. Finally, it is most desirable to use an unmodified, standard production CMOS process.

Low-power operation requires the entire system to operate on a 3 V supply. This supply voltage cannot be any lower because in many places the analog circuits contain stacked cascode transistors, and the FET threshold voltages are about 0.8 V. RF amplifiers are normally tuned with inductor loads, which on silicon IC's are almost always off-chip passive components [151]. However, this is itself wasteful of power, because in addition to the current supplied into the inductor, the circuit must also drive the various parasitic capacitances in off-chip routing the RF signal.



(a) (b)

Fig. 21. (a) Balanced input CMOS RF amplifier. Cascode stage driving on-chip inductor substantially enhances voltage gain. Outputs common-source stages are for measurement only. LNA will directly drive mixer. (b) A microphotograph of a 110 nH spiral inductor suspended over a pit on substrate.

Spiral inductors on silicon substrates suffer from a large capacitance to the conducting silicon substrate. However, after many years of the belief that no useful inductors could be made on silicon, it was found that spiral inductors as large as 10 nH with self-resonance beyond 2 GHz, could be fabricated with the standard interconnect metallization. These inductors were used to build passive filters, a tuned amplifier, and an LC voltage-controlled oscillator on a silicon bipolar IC [152]–[154]. This work has rekindled interest in the design and modelling of small value spiral inductors on silicon (usually bipolar) IC's, with values in the range of 1–10 nH [155], [156]. In some instances, higher inductor Q's are obtained with thicker metallization, or higher resonant frequencies with thicker oxides. In search of larger value inductors for low-power, high-gain amplifiers, we have developed a fabrication method whereby the silicon substrate is selectively removed under the inductor, substantially reducing the capacitance to the substrate, and thus extending self-resonance to a higher frequency (Fig. 21(b)) [157]. The self-resonant frequency is now almost as high as it is on a semi-insulating GaAs substrate. Using two such inductors as loads in a balanced circuit, a 900 MHz RF amplifier has been built in 1-μm CMOS with a 30 dB gain draining only 3 mA (Fig. 21(a)). The amplifier IP3 is about 0 dBm, owing to the wide linear range of the MOS differential pair.

The front-end mixer in a direct-conversion receiver must be highly linear to suppress unwanted intermodulation produced by interferers. Also, as the local oscillator frequency is centered at RF in this direct-conversion receiver, inadequate reverse isolation or shielding may cause this frequency to leak through the antenna and to interfere with other nearby receivers tuned to the same RF. We use an unusual mixer—one that happens to be particularly well suited to MOS implementation—to circumvent both these problems. The desired signal is downconverted to baseband by sub-sampling the incoming RF [158]. The mixer circuit is a track-and-hold, with such a wide track-mode bandwidth that it can follow the RF waveform, and a short enough aperture to acquire the *instantaneous* value of the RF waveform on receipt of the sampling clock edge (Fig. 22). The received RF in our system is a 26 MHz wide spread-spectrum centered on a 915 MHz carrier, so by sub-sampling this waveform at a clock rate of 52 MHz or higher, the spectrum is translated to baseband without aliasing. However, the mixer also acquires wideband noise accompanying the signal from dc up to the gigahertz track bandwidth, and aliases the rms noise into a bandwidth at half the sample-rate, thereby raising the baseband noise spectral density by the ratio of the track-bandwidth to the sample-rate. Sub-sampling mixers therefore have a higher noise figure than conventional mixers. However, as such a mixer tends to be more linear than an analog multiplier, its dynamic range is also large, provided a high-gain, low-noise RF amplifier precedes it. The overall front-end spurious-free dynamic range is jointly set by the individual specifications of the RF amplifier and mixer. In this 1 μm CMOS prototype, the mixer circuit draws 4 mA from 3-V to acquire samples of a 900 MHz modulated waveform at a 50 MHz rate, which it then translates to baseband. The linearity, as measured by a +26 dBm third-order intercept, exceeds that of most continuous-time 900 MHz monolithic mixers, and for the fundamental reasons given above, the noise figure is 18 dB.

A low-power, agile frequency synthesizer with high spectral purity is required in a frequency-hopped transceiver. Conventional synthesizers based on a phase-locked loop with a variable-modulus divider in feedback suffer from limited agility, because the loop bandwidth may limit the speed of frequency switching. It is also difficult to design a VCO which both covers a wide range of frequency spreading, and produces a pure spectrum. The DDFS-DAC combination is an alternative solution to this problem, which so far has only been used by the military in spread-spectrum communication systems. A direct-digital frequency synthesizer (DDFS) [159] consists of an accumu-

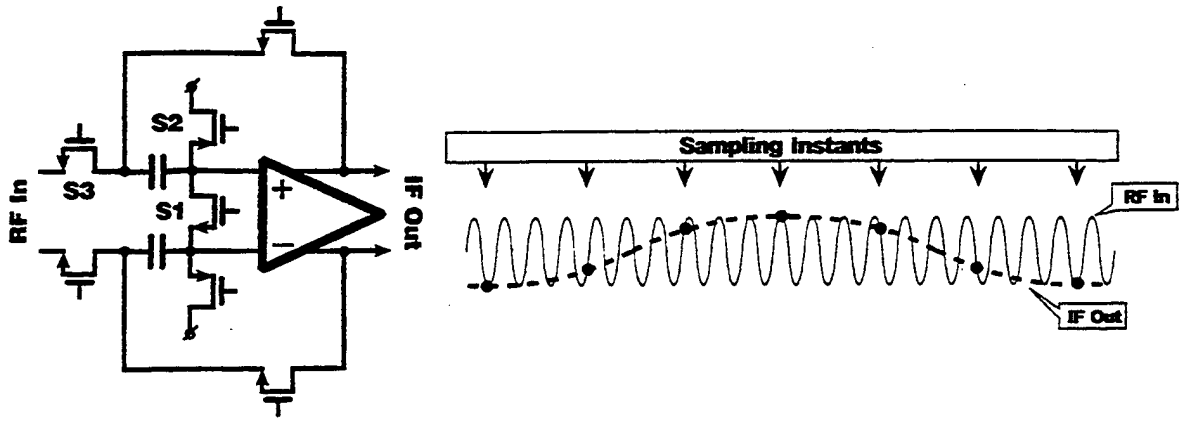
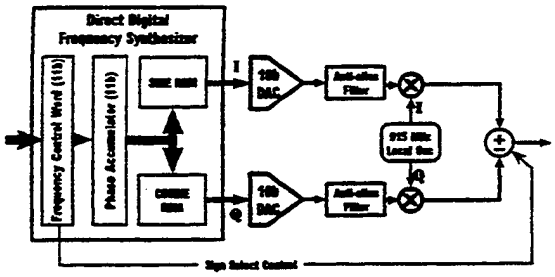


Fig. 22. A downconversion mixer based on direct sampling of RF. During track mode, switches are closed, and bandwidth of passive circuit is about 1 GHz. On receipt of sample clock, switches open within very short aperture time, capuring instantaneous value of RF waveform. Op amp removes switch charge injection, and buffers to subsequent circuits.

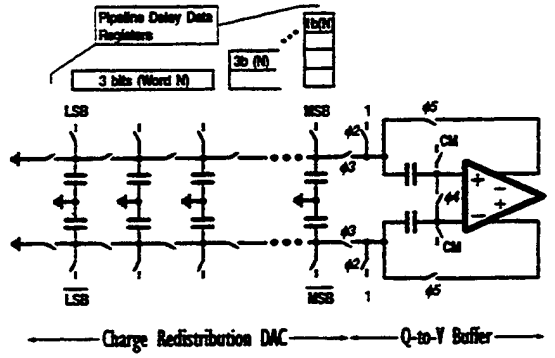
lator and ROM, which produces a sequence of digital words representing samples of a sinewave at a frequency set by an input control word. A D/A Converter converts the digital samples into a discrete-time analog sinewave, and a subsequent analog filter may smooth this into a continuous-time sinewave. Roundoff errors in the DDFS and nonlinearity in the DAC and filter will limit the attainable spectral purity of the synthesized sinewave [159]. In practice, DAC imperfections may contribute the largest nonlinearity in the system. A characteristic feature of digitally synthesized sinewaves is that the major imperfections do not necessarily produce harmonic tones, but instead spurious tones at rational multiples of the fundamental. These spurious tones often cluster around the fundamental and cannot be filtered. The DAC can add spurious tones of its own. We use a highly efficient DDFS combined with a high-speed, low-power charge-redistribution DAC to produce an FSK-modulated, frequency-hopped spread-spectrum at baseband, which a fixed-frequency local oscillator upconverts to RF. A CMOS implementation of a DDFS-DAC at 3 V dissipates only 40 mW when clocked at 50 MHz (of which the quadrature ROM-accumulator accounts for 35 mW), with worst-case spurious levels of −57 dBc or less across the entire spreading range (Fig. 23) [160].

The 915 MHz local oscillator is a four-stage MOS ring oscillator locked in a PLL to a lower frequency crystal reference. The quadrature outputs at 915 MHz for the image-reject mixers are tapped at diametrically opposite points [130], [134] to a phase accuracy of better than 1°. As the local-oscillator operates at a fixed frequency, it is locked to a low-frequency crystal reference in a PLL solely optimized to reduce phase noise. At a 100 kHz offset, a phase noise level of −75 dBc/Hz is measured on a prototype.

The power amplifier is a binary-weighted array of FET's biased near threshold, which is digitally selected to deliver power levels from −15 dBm to +15 dBm through a matching network to the antenna with a 40% conversion efficiency (Fig. 24) [161].



(a)



(b)

Fig. 23. (a) An agile frequency synthesizer. The DDFS produces digital samples of a sine and cosine, at a frequency set by an input word. DAC's convert these to analog domain, and after quadrature upconversion, either the upper or lower sideband is selected. Thus 0–13 MHz at baseband frequency produces output spanning 902–928 MHz. First-order anti-alias filter suffices if DDFS/DAC operates at 80 MHz. (b) DAC uses passive, pipelined architecture.

The transceiver achieves spatial diversity with two individual receive channels, which are powered-up continuously and connected to two different types of miniature antennas. Data decisions are made on an equally weighted sum of their baseband outputs. A high-order switched-capacitor elliptic lowpass filter selects the desired channel, which the on-board DDFS-based synthesizer has dehopped

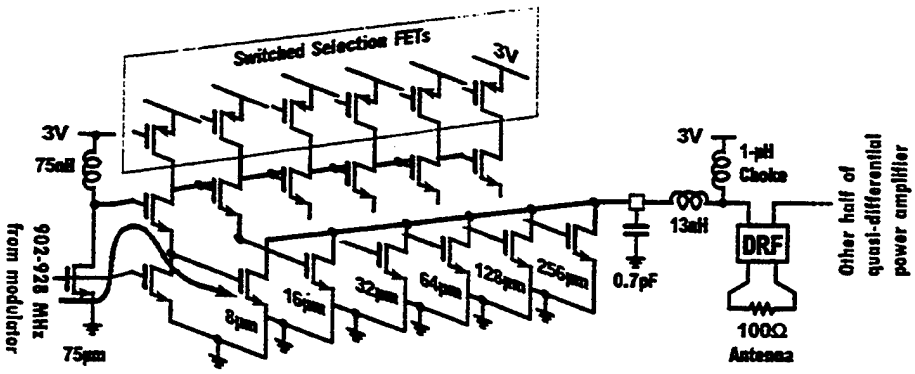


Fig. 24. A digitally controlled power amplifier. The binary-weighted array of output NMOSFET's drive the antenna load differentially, to deliver up to +15 dBm from a 3-V supply. Large on-chip inductor used as exciter load gives swings above power supply. PMOSFET's selectively activate source follower buffers, and thus total delivered power.

to dc in the course of downconversion. The capacitively coupled output is then amplitude-limited to produce a one-bit stream fed to an optimal FSK digital detector [39]. The detector quantizes the zero-crossings in time by oversampling the limited baseband data, and seeks correlations between one-bit representations of the two possible data values. A logarithmic amplifier connected to taps on the limiting amplifier chain measures the received signal-strength [162]. Operating from a 3-V supply, the 1-μm CMOS transceiver is expected to drain 70 mA in receive mode, and 100 mA in transmit mode (the power amplifier is on-chip). Considering that the transceiver communicates on a 26 MHz-wide spectrum, this is indeed a low power dissipation.

This research project is now exploring issues related to the single-chip integration of this transceiver. Such a highly integrated CMOS "VLSI radio" would represent a major step forward in the evolution of the radio integrated circuits described in this paper.

X. Conclusions

This survey has covered some of the key existing and emerging communication applications which have prompted advances in RF IC's. The emphasis was on widespread *portable consumer* applications. This has excluded coverage of integrated television tuners, for instance, which were historically the principal motivation for the development of UHF RF-IC's by the consumer electronics industry in the mid-1970's [163]. The fact that about 18 million tuners are produced every year means that there has been a steady stream of innovations in the building-block and architectures for this application [164]–[178]. Then there are emerging areas, such as personal GPS receivers, where RF integration and low-power will be the key for acceptance by consumers [179], [180]. Here, too, RF-IC designers are responding with integrated front ends [181]–[184], although the greater challenges to realize a system with an acceptable precision may lie in the baseband signal processing.

The lucrative global wireless market is attracting a great deal of attention from circuit designers across the industry. The RF range of interest spans 400 MHz to 2.5 GHz. There is a multi-pronged industry-wide assault to provide the right solutions. The solution must be low-cost, low-power, and it must give high functionality. MMIC designers bring to RF-IC's a microwave style derived from small-scale circuits which operated at much higher frequencies. On the other hand, as the frequency range lies within the capability of modern silicon IC technologies, there is a response from this community. Bipolar and BiCMOS RF-IC's are often extensions of baseband style circuits to high-frequency, often with more functionality than GaAs MMICs offer. A low transistor f_T is no longer a handicap, and experience rapidly accumulates on how to solve some of the unique problems of silicon substrates, such as high-frequency losses in the substrate [185] and on-chip coupling problems between circuit blocks. In the not too distant future, as microprocessors and memories drive the linewidths of silicon IC processes to deep submicron, RF-IC's might be designed as untuned wideband circuits, in the style of IC's for video applications today.

Small-scale GaAs MMIC's might sometimes perform better as standalone components, but embedded in a system such as a digital cellular telephone or a spread-spectrum transceiver, they require a large overhead in support circuits, and thereby lose an edge in certain specifications, such as power dissipation. Silicon RF-IC's offer higher levels of integration, but even within the silicon IC design community there are differing points of view. While integrated silicon bipolar transceivers are to be found in the next generation of products, CMOS RF solutions, albeit at an exploratory stage, now look very promising. They combine elements from well established techniques in voiceband and video IC's, and are unfettered in partitioning tasks between analog and digital signal processing, while using off-chip passive components when necessary. Digital control components and baseband signal processors are readily integrated with, indeed merged into, the RF and IF

sections. This array of competing technologies offers communication system designers more creative opportunity, and the best wireless transceiver solutions may well emerge from system design evolving together with architecture, circuits, antennas, and power-allocation plans.

What are the prospects for this multitude of approaches to low-power wireless communications? If we project on the experience gained from baseband and video telecommunications circuits, the future wireless transceiver may be only one BiCMOS or CMOS large-scale, mixed analog-digital IC, requiring one passive filter and a crystal resonator. Baseband digital signal processing may make up for imperfections in the front-end RF sections. As it will likely be used in a microcell environment at low transmit power levels, the transceiver may be encapsulated in a plastic package. There is no fundamental reason why it should need more than a single 1.5 V battery as the power source. After decades of thinking of it only as science-fiction, the electronics world is prodded forth by the communications explosion into making the two-way wristwatch-size radio a reality. A large community researchers, not trained in the classical radio art, is busy reexamining the conventions, and at time challenging the received wisdom. We are poised at a turning point in the evolution of radio-circuit engineering.

ACKNOWLEDGMENT

My own involvement in RF-IC design is considerably enriched by collaborations with the many talented colleagues and graduate students participating in the Low-Power Wireless Transceiver Program in the UCLA Electrical Engineering Department. Professors Henry Samueli, Greg Pottie, and Yahya Rahmat-Samii have made an exemplary teamwork possible on this project.

The recent financial stringencies at UC have not so far eroded the marvellously rich Engineering Collection and online facilities at the UCLA Science and Engineering Library. It would be impossible to write this paper without ready access to these resources. Karen Andrews and Aggi Raeder, two of the expert librarians there, enthusiastically helped me in using the latest tools to peruse the literature, and amazed me with the ease with which they could track down obscure references from all corners of the world.

I thank Prof. Bob Meyer for a very thorough reading of the draft manuscript and for his thoughtful comments on it, and also Taiwa Okanobu and Mark McDonald who made useful suggestions. I am very grateful to Dr. Ran-Hong Yan, the Guest Editor of the *Proceedings of the IEEE* who solicited this manuscript, for his extraordinary ability to obtain speedy reviews and patiently absorb delays on my part in bringing this work to completion.

REFERENCES

[1] G. H. Heilmeier, "Personal communications: *Quo Vadis*," in *Int. Solid State Circ. Conf.*, San Francisco, 1992, pp. 24–26.
[2] D. C. Cox, "Universal digital portable radio communications," *Proc. IEEE*, vol. 75, pp. 436–476, Apr. 1987.
[3] S. Sheng, A. Chandrakasan, and R. W. Brodersen, "A portable multimedia terminal," *IEEE Comm. Magazine*, vol. 30, no. 12, pp. 64–75, Dec. 1992.
[4] B. A. Beatty, "The IBM personal communicator—Design considerations," *SOUTHEASTCON*, Miami, FL, 1994, pp. 432–435.
[5] A. Matsuzawa, "Low-voltage and low-power circuit design for mixed analog/digital systems in portable equipment," *IEEE J. Solid-State Circ.*, vol. 29, pp. 470–480, Apr. 1994.
[6] E. A. Vittoz, "Low-power design: Ways to approach the limits," in *Int. Solid State Circ. Conf.*, San Francisco, 1994, pp. 14–18.
[7] F. E. Terman, *Radio Engineers' Handbook*. New York: McGraw-Hill, 1943.
[8] W. L. Everitt, Ed., *Fundamentals of Radio and Electronics*, 2nd ed. Englewood Cliffs, NJ: Prentice-Hall, 1958.
[9] K. K. Clarke and D. T. Hess, *Communication Circuits: Analysis and Design*. Reading, MA: Addison-Wesley, 1971.
[10] H. L. Krauss, C. W. Bostian, and F. H. Raab, *Solid State Radio Engineering*. New York: Wiley, 1980.
[11] W. H. Hayward, *Introduction to Radio Frequency Design*. Englewood Cliffs, NJ: Prentice-Hall, 1982.
[12] U. L. Rohde and T. T. N. Bucher, *Communications Receivers: Principles & Design*. New York: McGraw-Hill, 1988.
[13] D. O. Pederson and K. Mayaram, *Analog Integrated Circuits for Communications: Principles, Simulation, and Design*. Boston: Kluwer, 1991.
[14] M. Ibuka, "How Sony developed electronics for the world market," *IEEE Trans. Eng. Manage.*, vol. EM-22, pp. 15–19, Feb. 1975.
[15] A. Laundrie and K. Negus, "GPS receiver chip targets low-power commercial markets," in *Microwaves & RF*, Aug. 1992, pp. 135–138.
[16] N. Camilleri, D. Lovelace, J. Costa, and D. Ngo, "New development trends for silicon RF device technologies," in *Microwave & Millimeter-Wave Monolithic Circ. Symp.*, San Diego, CA, 1994, pp. 5–8.
[17] W. Kanow and I. Siewert, "Integrated circuits for hi-fi radios and tuners," in *Electronic Components and Applications*, vol. 4, no. 1, pp. 11–27, Nov. 1981.
[18] R. M. Warner and J. N. Fordemwalt, Eds., *Integrated Circuits: Design Principles and Fabrication*. New York: McGraw-Hill, 1965.
[19] C. S. Meyer, D. K. Lynn, and D. J. Hamilton, Eds., *Analysis and Design of Integrated Circuits*. New York: McGraw-Hill, 1968.
[20] Y. Hishida and M. Kazama, "Mobile communication system: Its semiconductor technology," in *Hitachi Review*, vol. 42, no. 3, pp. 95–100, June 1993.
[21] B. Crouch, Ed., *Dick Tracy: America's Most Famous Detective*. Secaucus, NJ: Citadel, 1987.
[22] L. Blaser and T. Taira, "An am/fm radio subsystem IC," *IEEE Trans. Consumer Electron.*, vol. CE-23, pp. 129–136, Feb. 1977.
[23] W. Peil and R. J. McFayden, "A single chip am/fm integrated circuit radio," *IEEE Trans. Consumer Electron.*, vol. CE-23, pp. 424–431, Mar. 1977.
[24] O. L. Richards, "A complete am/fm signal processing system," *IEEE Trans. Consumer Electron.*, vol. CE-24, pp. 34–38, Jan. 1978.
[25] W. G. Kasperkovitz, "An integrated fm receiver," *Microelectronics Reliability*, vol. 21, no. 2, pp. 183–189, 1981.
[26] A. A. Abidi, "Noise in active resonators and the available dynamic range," *IEEE Trans. Circ. and Syst.*, vol. 39, pp. 296–299, Apr. 1992.
[27] W. G. Kasperkovitz, "Fm receivers for mono and stereo on a single chip," *Philips Tech. Rev.*, vol. 41, no. 6, pp. 169–182, 1983.
[28] W. H. A. van Dooremolen and M. Hufschmidt, "A complete fm radio on a chip," in *Electronic Components and Applications*, vol. 5, no. 3, pp. 159–170, 1983.
[29] J. Brilka and G. Sieboerger, "A single-chip fm radio," in *Int. Conf. on Consumer Electron.*, pp. 156–157, 1984.
[30] T. Okanobu, T. Tsuchiya, K. Abe, and Y. Ueki, "A complete single-chip am/fm radio integrated circuit," *IEEE Trans. Consumer Electron.*, vol. CE-28, pp. 393–408, Mar. 1982.
[31] T. Okanobu, H. Tomiyama, and H. Arimoto, "Advanced low-voltage single chip radio IC," *IEEE Trans. Consumer Electron.*, vol. 38, pp. 465–475, Mar. 1992.

[32] D. G. Tucker, "The history of the homodyne and synchrodyne," *J. Brit. Inst. of Radio Eng.*, vol. 14, no. 4, pp. 143–154, 1954.
[33] D. Mitchell and K. G. van Wynen, "A 150-mc personal radio signaling system," *Bell Syst. Tech. J.*, vol. XL, no. 5, pp. 1239–1257, 1961.
[34] A. E. Kerwien and L. H. Steiff, "Design of a 150-megacycle pocket receiver for the BELLBOY personal signaling system," *Bell Syst. Tech. J.*, vol. XLII, no. 3, pp. 527–565, 1963.
[35] K. Nagata, M. Akahori, T. Mori, and S. Umetsu, "Digital display radio paging system," in *NEC Res. & Devel.*, no. 68, pp. 16–23, Jan. 1983.
[36] I. A. W: Vance, "An integrated circuit VHF radio receiver," *The Radio and Electronic Engineer*, vol. 50, no. 4, pp. 158–164, 1980.
[37] ——, "Fully integrated radio paging receiver," *IEE Proc.*, vol. 129, pt. F, no. 1, pp. 2–6, 1982.
[38] R. C. French, "A high technology VHF radio paging receiver," in *Int. Conf. on Mobile Radio Syst. and Tech.*, York, UK, 1984, pp. 11–15.
[39] J. Min, A. Rofougaran, H. Samueli, and A. A. Abidi, "An all-CMOS architecture for a low-power frequency-hopped 900 MHz spread-spectrum transceiver," in *Custom IC Conf.*, San Diego, CA, 1994, pp. 379–382.
[40] K. Holtvoeth, "A fully integrated low-power FSK receiver for VHF and UHF paging," *IEEE Trans. Consumer Electron.*, vol. 35, pp. 707–713, Mar. 1989.
[41] S. Drude and T. Rudolph, "New chip-set reduces power consumption of radio pagers," *Electron. Components and Applications*, vol. 9, no. 2, pp. 112–124, 1989.
[42] J. F. Wilson, R. Youell, T. H. Richards, G. Luff, and R. Pilaski, "A single-chip VHF and UHF receiver for radio paging," *IEEE J. Solid State Circ.*, vol. 26, pp. 1944–1950, Dec. 1991.
[43] K. Nagata, D. Ishii, T. Mori, T. Oyagi, and T. Seo, "Slim digital pagers, in *NEC Res. & Develop.*, no. 74, pp. 55–64, July 1984.
[44] K. Yamasaki, S. Yoshizawa, Y. Minami, T. Asai, Y. Nakano, and M. Kuroda, "Compact size numeric display pager with new receiving system," in *NEC Res. and Develop.*, vol. 33, no. 1, pp. 73–81, Jan. 1992.
[45] A. Burt, "Direct conversion receivers come of age in the paging world," in *GEC Rev.*, vol. 7, no. 3, pp. 156–160, 1992.
[46] S. Tanaka, A. Nakajima, J. Nakagawa, and A. Nakagoshi, "High-frequency, low-voltage circuit technology for VHF paging receiver," *IEICE Trans. Fundamentals of Electron., Commun. and Comp. Sci.*, vol. E76-A, no. 2, pp. 156–163, 1993.
[47] B. F. Cambou, "MMIC consumer application and production," in *Microwave & Millimeter-Wave Monolithic Circ. Symp.*, Albuquerque, NM, 1992, pp. 1–3.
[48] H. Komatsu, R. Norose, and S. J. Symonds, "The wristwatch as a personal communications device," in *Pacific Telecommun. Conf.*, Honolulu, 1993, pp. 242–244.
[49] M. Pardoen, R. Pache, and E. Dijkstra, "Direct-conversion receiver provides CP-FSK operation," in *Microwaves & RF*, Mar. 1994, pp. 151–155.
[50] G. J. Lomer, "Telephoning on the move—Dick Tracy to Captain Kirk," *IEE Proc.*, vol. 134, pt. F, no. 1, pp. 1–8, 1987.
[51] K. Wakino, "Recent developments of dielectric resonator materials and filters in Japan," *Ferroelectrics*, vol. 91, pp. 69–86, 1989.
[52] M. Hikita, T. Tabuchi, and A. Sumioka, "Miniaturized SAW devices for radio communication transceivers," *IEEE Trans. Vehic. Tech.*, vol. 38, pp. 2–8, Jan. 1989.
[53] K. Eda, "Ultra-small SAW filter works in mobile communications," in *J. Electron. Eng.*, pp. 54–57, May 1992.
[54] Y. Tamura, T. Maru, N. Hirasawa, H. Okuno, M. Komoda, and M. Hotsumi, "Hand-held portable equipment for cellular mobile telephone, in *NEC Res. and Devel.*, no. 87, pp. 34–43, Oct. 1987.
[55] D. Anderson and R. J. Zavrel, "RF ICs for portable communications equipment," in *Electron. Components and Appl.*, vol. 7, no. 1 pp. 37–44, 1985.
[56] I. A. Koullias, S. L. Forgues, and P. C. Davis, "A 100 MHz IF amplifier/quadrature demodulator for GSM cellular radio mobile terminals," in *Bipolar Circ. and Tech. Mtg.*, Minneapolis, MN, 1990, pp. 248–251.
[57] Y. Mori, H. Yabuki, M. Ohba, M. Sagawa, M. Makimoto, and I. Shibazaki, "Miniaturized RF-circuit modules for land mobile communication equipment," in *IEEE Vehic. Tech. Conf.*, Orlando, FL, 1990, pp. 65–70.
[58] M. Hikita, A. Yuhara, and K. Oda, "New low-loss, miniature SAW filters," in *Hitachi Rev.*, vol. 42, no. 3, pp. 119–124, June 1993.
[59] L. Scharf, H. Frank, and T. Pollakowski, "High f_T and low noise," in *Siemens Components*, vol. 25, no. 4, pp. 119–122, 1990.
[60] S. Yamada and K. Oguri, "Low-power front-end oriented discrete semiconductor devices," in *Hitachi Rev.*, vol. 42, no. 3, pp. 101–106, June 1993.
[61] C. Huang, "Where do integrated circuits replace discretes?," in *Applied Microwaves & Wireless*, pp. 5–6, Summer 1994.
[62] S. P. MacCabe, "GaAs MMICs for cellular telephones, GPS and DBS markets," in *WESCON*, San Francisco, 1989, pp. 279–281.
[63] M. Rocchi, "GaAs & Si MMIC building blocks: A moot point revisited," *Microelectron. Eng.*, vol. 15, pp. 685–692, 1991.
[64] Y. Aoki, K. Kobayashi, and W. Kennan, "Monolithic microwave ICs (MMICs) for telecommunication," in *Fujitsu Sci. & Tech. J.*, vol. 30, no. 1, pp. 32–39, June 1994.
[65] Y. Imai, M. Tokumitsu, A. Minakawa, T. Sugeta, and M. Aikawa, "Very low-current and small size GaAs MMICs for L-band front-end applications," in *GaAs IC Symp.*, Monterey, CA, 1989, pp. 71–74.
[66] Y. Imai, M. Tokumitsu, and A. Minakawa, "Design and performance of low-current GaAs MMICs for L-band front-end applications," *IEEE Trans. Microwave Theory and Tech.*, vol. 39, pp. 209–215, Feb. 1991.
[67] K. R. Cioffi, "Ultra-low dc power consumptions in monolithic L-band components," *IEEE Trans. Microwave Theory & Techn.*, vol. 40, pp. 2467–2472, Dec. 1992.
[68] C. Kermarrec, M. Takano, F. McGrath, P. O'Sullivan, G. Dawe, E. Heaney, and G. St. Onge, "High performance, low cost GaAs MMICs for personal phone applications," in *Int. Symp. on GaAs and Related Compounds*, Karuizawa, Japan, 1992.
[69] E. Heaney, F. McGrath, P. O'Sullivan, and C. Kermarrec, "Ultra low power low noise amplifiers for wireless communications," in *GaAs IC Symp.*, San Jose, CA, 1993, pp. 49–51.
[70] S. Hara, K. Osato, A. Yamada, T. Tsukao, and T. Yoshimasu, "Miniaturized low noise variable MMIC amplifiers with low power consumption for L-band portable communication applications," in *Microwave and Millimeter-Wave Monolithic Circ. Symp.*, Atlanta, GA, 1993, pp. 67–70.
[71] T. Ohgihara, S. Kusunoki, M. Wada, and Y. Murakami, "GaAs JFET front-end MMICs for L-band personal communications," in *Microwave and Millimeter-Wave Monolithic Circ. Symp.*, Atlanta, GA, 1993, pp. 9–12.
[72] C. Kusano, E. Hase, and K. Sakamoto, "Low-power, low-noise GaAs MMIC," in *Hitachi Rev.*, vol. 42, no. 3, pp. 107–112, June 1993.
[73] R. E. Lehmann and D. D. Heston, "X-band monolithic series feedback LNA," *IEEE Trans. Microwave Theory and Tech.*, vol. MTT-33, pp. 1560–1566, Dec. 1985.
[74] H. Sakai, A. Tezuka, Y. Mori, M. Sagawa, T. Katoh, J. Itoh, and K. Fujimoto, "Low-power GaAs ICs for mobile communication equipment," in *Int. Symp. on GaAs and Related Compounds*, Karuizawa, Japan, 1992.
[75] V. Nair, R. Vaitkus, D. Scheitlin, J. Kline, and H. Swanson, "Low current GaAs integrated down converter for portable communication applications," in *GaAs IC Symp.*, San Jose, CA, 1993, pp. 41–44.
[76] C. Woo, A. Podell, R. Benton, D. Fisher, and J. Wachsman, "A fully integrated transceiver chip for the 900 MHz communication bands," in *GaAs IC Symp.*, Miami Beach, FL, 1992, pp. 143–146.
[77] P. E. Chadwick, "High performance integrated circuit mixers," in *Radio Receivers & Assoc. Syst.*, London, 1981, pp. 1–9.
[78] D. M. Snider, "A theoretical analysis and experimental confirmation of the optimally loaded and overdriven RF power amplifier," *IEEE Trans. Electron Devices*, vol. ED-14, pp. 851–857, Dec. 1967.
[79] J. J. Komiak, "Design and performance of an octave band 11 watt power amplifier MMIC," *IEEE Trans. Microwave Theory and Techn.*, vol. 38, pp. 2001–2006, Dec. 1990.
[80] T. Takagi, Y. Ikeda, K. Seino, G. Toyoshima, and A. Inoue, "A UHF band 1.3 W monolithic amplifier with efficiency of 63%," in *Microwave and Millimeter-Wave Monolithic Circ. Symp.*, Albuquerque, NM, 1992, pp. 35–38.

[81] M. Gat, D. S. Day, S. Chan, C. Hua, and J. R. Basset, "A 3.0 watt high efficiency C-band power MMIC," in *GaAs IC Symp.*, Monterey, CA, 1991, pp. 331–334.
[82] O. Ishikawa, M. Maeda, K. Nishii, M. Yanagihara , T. Yokoyama, Y. Ikeda, and Y. Ota, "Cellular telecommunication GaAs power modules," in *Applied Microwave and Wireless*, vol. 4, no. 3, pp. 83–88, Fall 1992.
[83] M. Maeda, M. Nishijima, H. Takehara, C. Adachi, H. Fujimoto, and O. Ishikawa, "A 3.5 V, 1.3 W GaAs power multi-chip IC for cellular phones," *IEEE J. Solid-State Circ.*, vol. 29, pp. 1250–1256, Oct. 1994.
[84] I. Yoshida, M. Katsueda, S. Ohtaka, Y. Maruyama, and T. Okabe, "Highly efficient 1.5 GHz Si power MOSFET for digital cellular front end," in *Int. Symp. on Power Semicond. Devices*, Tokyo, 1992, pp. 156–157.
[85] T. Okabe and K. Kobayashi, "Radio-frequency MOS power module," in *Hitachi Rev.*, vol. 42, no. 3, pp. 113–118, June 1993.
[86] R. G. Meyer and W. D. Mack, "A 1 GHz BiCMOS RF front-end IC," *IEEE J. Solid-State Circ.*, vol. 29, no. 3, pp. 350–355, 1994.
[87] M. N. Sera, "Chip set addresses North American digital cellular market," in *RF Design*, vol. 17, no. 3, pp. 54–62, Mar. 1994.
[88] Y. Funada, "VCO techniques enable compact, high-frequency mobilecom," in *J. Electron. Eng.*, Dec. 1993, pp. 40–43.
[89] Y. Tamura, A. Yonehata, S. Miyazaki, F. Kobayasi, Y. Fukuda, and J. Kamishiro, "Development of advanced mobile telephone P3 (personal pocket phone)," in *NEC Res. and Develop.*, no. 98, pp. 60-70, July 1990.
[90] H. Ohmori, "SAW filters hold key to creating small, portable equipment," in *J. Electron. Eng.*, pp. 63–67, May 1992.
[91] G. Kainz, "Microwave ceramics boosts miniaturization," in *Siemens Components*, vol. 27, no. 5, pp. 11–14, Sept./Oct. 1992.
[92] Y. Kado, M. Suzuki, K. Koike, and Y. Omura, "A 1 GHz/0.9 mW CMOS/SIMOX divide-by-128/129 dual-modulus prescaler using a divide-by-2/3 synchronous counter," *IEEE J. Solid-State Circ.*, vol. 28, pp. 513–517, Apr. 1993.
[93] T. Seneff, L. McKay, K. Sakamoto, and N. Tracht, "A Sub-1 mA 1.5 GHz silicon bipolar dual modulus prescaler," *IEEE J. Solid-State Circ.*, vol. 29, pp. 1206–1211, Oct. 1994.
[94] M. Morishima, "Present and future of mobile communications in Japan," in *Microwave & Millimeter-Wave Monolithic Circ. Conf.*, Atlanta, GA, 1993, pp. 3–6.
[95] I. Shimizu, S. Urabe, K. Hirade, K. Nagata, and S. Yuki, "A new pocket-size cellular telephone for NTT high-capacity land mobile communication system," in *Vehic. Tech. Conf.*, St. Louis, MO, 1991, pp. 114–119.
[96] I. A. Koullias, J. H. Havens, I. G. Post, and P. E. Bronner, "A 900 MHz transceiver chip set for dual-mode cellular radio terminals," in *Int. Solid State Circ. Conf.*, San Francisco, 1993, pp. 140–141.
[97] L. Longo, R. Halim, B.-R. Horng, K. Hsu, and D. Shamlou, "A cellular analog front end with a 98 dB IF receiver," in *Int. Solid State Circ. Conf.*, San Francisco, 1994, pp. 36–37.
[98] D. J. Goodman, "Second generation wireless information networks," *IEEE Trans. Vehic. Tech.*, vol. 40, pp. 366–374, Mar. 1991.
[99] J. Fenk, W. Birth, R. G. Irvine, P. Sehrig, and K. R. Schon, "An RF front-end for digital mobile radio," in *Bipolar Circ. and Tech. Mtg.*, Minneapolis, MN, 1990, pp. 244–247.
[100] E. Badura, "New integrated circuits for GSM mobile radio," in *Siemens Components*, vol. 26, no. 2, pp. 72–76, 1991.
[101] S. Beyer and G. Lipperer, "Low-current oscillator design for 900 MHz GSM applications," in *Microwave Eng. Europe*, Oct. 1991, pp. 35–41.
[102] R. E. Jenkins, "Transceiver chip simplifies GSM cellular design," in *RF Design*, June 1994, pp. 26–35.
[103] J. Sevenhans, A. Vanwelsenaers, J. Wenin, and J. Baro, "An integrated Si bipolar RF transceiver for a zero IF 900 MHz GSM digital radio front-end of a hand portable phone," in *Custom IC Conf.*, San Diego, CA, 1991, pp. 7.7/1–4.
[104] J. Sevenhans and A. Vanwelsenaers, "A silicon transceiver for a 900 MHz GSM hand set," in *Microwave Eng. Europe*, Dec./Jan. 1992/1993, pp. 59–63.
[105] D. Haspeslagh, J. Ceuterick, L. Kiss, and J. Wenin, "BBTRX: A baseband transceiver for a zero IF GSM hand portable station," in *Custom IC Conf.*, San Diego, CA, 1992, pp. 10.7.1–10.7.4.
[106] M. McDonald, "A DECT transceiver chip set," in *Custom IC Conf.*, San Diego, CA, 1992, pp. 10.6.1–10.6.4.
[107] B. Madsen and D. E. Fague, "Radios for the future: Designing for DECT," in *RF Design*, Apr. 1993, pp. 48–53.
[108] M. Eccles, "Gigahertz systems on a chip," in *Electron. World+Wireless World*, Aug. 1993, pp. 662–663.
[109] I. White, "UHF technology for the cordless revolution," in *Electron. World+Wireless World*, July/Aug. 1993, pp. 542–546 and 657–661.
[110] A. Boveda, G. L. Bonato, and O. Ripolles, "GaAs monolithic 1.5–2.5 GHz image rejection receiver," in *Int. Symp. on Circ. & Systems*, San Diego, CA, 1992, pp. 232–235.
[111] A. Boveda, F. Ortigoso, and J. I. Alonso, "A 0.7–3 GHz GaAs QPSK/QAM direct modulator," in *Int. Solid State Circ. Conf.*, San Francisco, 1993, pp. 142–143.
[112] M. Williams, F. Bonn, C. Gong, and T. Quach, "GaAs RF ICs target 2.4 GHz frequency band," in *Microwaves & RF*, July 1994, pp. 111–118.
[113] M. D. McDonald, "A 2.5 GHz BiCMOS image reject front-end," in *Int. Solid State Circ. Conf.*, San Francisco, 1993, pp. 144–145.
[114] J. L. Wang, T. Tsuchiya, H. Takeuchi, K. Shirotori, S. Miyazaki, and T. Nakata, "A miniature ultra low-power 2. 5 GHz down-converter IC for wireless communications," in *NEC Res. and Develop.*, vol. 35, no. 1, pp. 46–50, Jan. 1994.
[115] T. Tsukahara, M. Ishikawa, and M. Muraguchi, "A 2 V 2GHz direct-conversion quadrature modulator," in *Int. Solid State Circ. Conf.*, San Francisco, 1994, pp. 40–41.
[116] M. Nakatsugawa, Y. Yamaguchi, and M. Muraguchi, "An L-band ultra low power consumption monolithic low noise amplifier," in *GaAs IC Symp.*, San Jose, CA, 1993, pp. 45–48.
[117] D. Ngo, B. Beckwith, P. O'Neil, and N. Camilleri, "Low voltage GaAs power amplifiers for personal communications at 1.9 GHz," in *MTT-S Int. Microwave Symp.*, Atlanta, GA, 1993, pp. 1461–1464.
[118] T. Kunihisa, T. Yokoyama, H. Fujimoto, K. Ishida, H. Takehara, and O. Ishikawa, "High efficiency, low adjacent channel leakage GaAs power MMIC for digital cordless telephone," in *Microwave and Millimeter-Wave Monolithic Circ. Symp.*, San Diego, CA, 1994, pp. 55–58.
[119] T. Tokumitsu, I. Toyoda, and M. Aikawa, "Low voltage, high power T/R switch MMIC using LC resonators," in *Microwave and Millimeter-Wave Monolithic Circ. Symp.*, Atlanta, GA, 1993, pp. 27–30.
[120] K. Miyatsuji, S. Nagata, N. Yoshikawa, K. Miyanaga, Y. Ohishi, and D. Ueda, "A GaAs high-power RF single-pole double-throw switch IC for digital mobile communication system," in *Int. Solid State Circ. Conf.*, San Francisco, 1994, pp. 34–35.
[121] P. O'Sullivan, G. St. Onge, E. Heaney, F. McGrath, and C. Kermarrec, "High performance integrated PA, T/R switch for 1. 9 GHz personal communications handsets," in *GaAs IC Symp.*, San Jose, CA, 1993, pp. 33–35.
[122] A. Bateman, D. M. Haines, and R. J. Wilkinson, "Linear transceiver architectures," in *IEEE Vehic. Tech. Conf.*, Philadelphia, 1988, pp. 478–484.
[123] G. Schultes, A. L. Scholtz, E. Bonek, and P. Veith, "A new incoherent direct conversion receiver," in *IEEE Vehicle Tech. Conf.*, Orlando, FL, 1990, pp. 668–674.
[124] G. Schultes, E. Bonek, P. Weger, and W. Herzog, "Basic performance of a direct conversion DECT receiver," *Electron. Lett.*, vol. 26, no. 21, pp. 1746–1748, 1990.
[125] J. Sevenhans, D. Haspeslagh, A. Delarbre, L. Kiss, Z. Chang, and J. F. Kukielka, "An analog radio front-end chip set for a 1.9 GHz mobile radio telephone application," in *Int. Solid State Circ. Conf.*, San Francisco, 1994, pp. 44–45.
[126] V. Thomas, J. Fenk, and S. Beyer, "A one-chip 2 GHz single superhet receiver for 2 Mb/s FSK radio communication," in *Int. Solid State Circ. Conf.*, San Francisco, 1994, pp. 42–43.
[127] W. Veit, J. Fenk, S. Ganser, K. Hadjizada, S. Heinen, H. Herrmann, and P. Sehrig, "A 2.7 V 800 MHz-2.1 GHz transceiver chipset for mobile radio applications in 25 GHz f_T Si-bipolar," in *Bipolar Circ. and Tech. Mtg.*, Minneapolis, 1994, pp. 175–178.
[128] D. K. Weaver, "A third method of generation and detection of single-sideband signals," *Proc. IRE*, vol. 44, no. 12, pp. 1703–1705, 1956.
[129] K. Negus, B. Koupal, J. Wholey, K. Carter, D. Millicker, C. Snapp, and N. Marion, "Highly integrated transmitter RFIC with

monolithic narrowband tuning for digital cellular handsets," in *Int. Solid State Circ. Conf.*, San Francisco, 1994, pp. 38–39.
[130] A. A. Abidi, "Radio-frequency integrated circuits for portable communications," in *Custom IC Conf.*, San Diego, CA, 1994, pp. 151–158.
[131] J. Wenin, "ICs for digital cellular communication," in *Europe. Solid State Circ. Conf.*, Ulm, Germany, 1994, pp. 1–10.
[132] H. Tsurumi and T. Maeda, "Design study on a direct conversion receiver front-end for 280 MHz, 900 MHz, and 2.6 GHz band radio communication systems," in *IEEE Vehic. Tech. Conf.*, St. Louis, MO, 1991, pp. 457–462.
[133] A. Bateman and D. M. Haines, "Direct conversion transceiver design for compact low-cost portable mobile radio terminals," in *IEEE Vehic. Tech. Conf.*, San Francisco, 1989, pp. 57–62.
[134] A. W. Buchwald and K. W. Martin, "High-speed voltage-controlled oscillator with quadrature outputs," *Electron. Lett.*, vol. 27, no. 4, pp. 309–310, 1991.
[135] S. K. Enam and A. A. Abidi, "NMOS ICs for clock and data regeneration in Gb/s optical-fiber receivers," *IEEE J. Solid State Circ.*, vol. 27, pp. 1763–1774, Dec. 1992.
[136] J. K. Cavers and M. W. Liao, "Adaptive compensation for imbalance and offset losses in direct conversion transceivers," *IEEE Trans. Vehic. Tech.*, vol. 42, no. 4, pp. 581–588, 1993.
[137] J. T. Taylor and J. K. Omura, "Spread spectrum technology: A solution to the personal communications services frequency allocation dilemma," in *IEEE Commun. Magazine*, vol. 29, no. 2, pp. 48–51, Feb. 1991.
[138] R. A. Scholtz, "The origins of spread-spectrum communications," *IEEE Trans. Commun.*, vol. COM-30, pp. 822–854, Oct. 1982.
[139] D. T. Magill, F. D. Natali, and G. P. Edwards, "Spread-spectrum technology for commercial applications," *Proc. IEEE*, vol. 82, pp. 572–584, Apr. 1994.
[140] R. C. Dixon, *Spread Spectrum Systems.* New York: Wiley, 1976.
[141] M. Leonard, "Wireless data links broaden LAN options," in *Electronic Design*, pp. 51–58, Mar. 1992.
[142] D. T. Magill, "A fully-integrated, digital, direct-sequence, spread spectrum modem ASIC," in *Int. Symp. on Personal, Indoor and Mobile Radio Commun.*, Boston, 1992, pp. 42–46.
[143] R. Jain, H. Samueli, P. T. Yang, and C. Chien, "Computer-aided design of a BPSK spread-spectrum chip set," *IEEE J. Solid-State Circ.*, vol. 27, pp. 44–58, Jan. 1992.
[144] C. Chien, P. Yang, E. Cohen, R. Jain, and H. Samueli, "A 12.7 Mchip/s all-digital BPSK direct sequence spread-spectrum IF transceiver in 1.2 μm CMOS," in *Int. Solid-State Circ. Conf.*, San Francisco, 1994, pp. 30–31.
[145] M. Leonard, "PCMCIA-sized radio links portable WLAN terminals," in *Electron. Design*, pp. 45–50, Aug. 1993.
[146] L. M. Devlin, B. J. Buck, J. C. Clifton, A. W. Dearn, and A. P. Long, "A 2.4 GHz single chip transceiver," in *Microwave & Millimeter-Wave Monolithic Circ. Symp.*, Atlanta, GA, 1993, pp. 23–26.
[147] T. Apel, E. Creviston, S. Ludvik, L. Quist, and B. Tuch, "A GaAs MMIC transceiver for 2.45 GHz wireless commercial products," in *Microwave & Millimeter-Wave Monolithic Circ. Symp.*, San Diego, CA, 1994, pp. 15–18.
[148] B. Khabbaz, A. Douglas, J. DeAngelis, L. Hongsmatip, V. Pellicia, W. Fahey, and G. Dawe, "A high performance 2.4 GHz transceiver chip set for high volume commercial applications," in *Microwave & Millimeter-Wave Monolithic Circ. Symp.*, San Diego, CA, 1994, pp. 11–14.
[149] M. S. Wang, M. Carriere, P. O'Sullivan, and B. Maoz, "A single-chip MMIC transceiver for 2.4 GHz spread spectrum communication," in *Microwave & Millimeter-Wave Monolithic Circuits Symp.*, San Diego, CA, 1994, pp. 19–22.
[150] J. Min, A. Rofougaran, V. Lin, M. Jensen, H. Samueli, A. A. Abidi, G. Pottie, and Y. Rahmat-Samii, "A low-power hand-held frequency-hopped spread spectrum transceiver hardware architecture," in *Virginia Tech's 3rd Symp. on Wireless Personal Commun.*, Blacksburg, VA, 1993, pp. 10/1–8.
[151] M. Sakakura and S. Skiest, "Ultra-miniature chip inductors serve at high frequency," in *J. Electron. Eng.*, pp. 48–51, Dec. 1993.
[152] N. M. Nguyen and R. G. Meyer, "Si IC-compatible inductors and LC passive filters," *IEEE J. Solid-State Circ.*, vol. 25, pp. 1028–1031, Apr. 1990.
[153] ——, "A silicon bipolar monolithic RF bandpass amplifier," *IEEE J. Solid-State Circ.*, vol. 27, pp. 123–127, Jan. 1992.
[154] N. M. Nguyen and R. G. Meyer, "A 1.8 GHz monolithic LC voltage-controlled oscillator," *IEEE J. Solid-State Circ.*, vol. 27, pp. 444–450, Mar. 1992.
[155] D. Lovelace, N. Camilleri, and G. Kannell, "Silicon MMIC inductor modeling for high volume, low cost applications," in *Microwave J.*, vol. 37, no. 8, pp. 60–71, Aug. 1994.
[156] K. B. Ashby, W. C. Finley, J. J. Bastek, S. Moinian, and I. A. Koullias, "High Q inductors for wireless applications in a complementary silicon bipolar process," in *Bipolar Circuits and Tech. Mtg.*, Minneapolis, 1994, pp. 179–182.
[157] J. Y.-C. Chang, A. A. Abidi, and M. Gaitan, "Large suspended inductors on silicon and their use in a 2 μm CMOS RF amplifier," *IEEE Electron Device Lett.*, vol. 14, pp. 246–248, May 1993.
[158] P. Y. Chan, A. Rofougaran, K. A. Ahmed, and A. A. Abidi, "A highly linear 1 GHz CMOS downconversion mixer," in *Europe. Solid-State Circ. Conf.*, Sevilla, Spain, 1993, pp. 210–213.
[159] H. T. Nicholas and H. Samueli, "A 150 MHz direct digital frequency synthesizer in 1.25 μm CMOS with −90 dBc spurious response," *IEEE J. Solid-State Circ.*, vol. 26, pp. 1959–1969, 1991.
[160] G. Chang, A. Rofougaran, M. K. Ku, A. A. Abidi, and H. Samueli, "A low-power CMOS digitally synthesized 0–13 MHz agile sinewave generator," in *Int. Solid-State Circ. Conf.*, San Francisco, 1994, pp. 32–33.
[161] M. Rofougaran, A. Rofougaran, C. Olgaard, and A. A. Abidi, "A 900 MHz CMOS RF power amplifier with programmable output," in *Symp. on VLSI Circ.*, Honolulu, 1994, pp. 133–134.
[162] K. Kimura, "A CMOS logarithmic amplifier with unbalanced source-coupled pairs," *IEEE J. Solid-State Circ.*, vol. 28, pp. 78–83, Jan. 1993.
[163] S. Watanabe, "Technology transfer for high frequency devices for consumer electronics," *IEEE Trans. Microwave Theory and Tech.*, vol. 40, pp. 2461–2466, Dec. 1992.
[164] U. Ablassmeier, W. Kellner, and H. Kniepkamp, "GaAs FET upconverter for TV tuner," *IEEE Trans. Electron Devices*, vol. ED-27, pp. 1156–1159, June 1980.
[165] K. Torii, S. Fujimori, S. Komatsu, S. Shimizu, K. Yoshihasa, J. Nagai, and Y. Yamagata, "Monolithic integrated VHF TV tuner," *IEEE Trans. Consum. Electron.*, vol. CE-26, pp. 180–187, 1980.
[166] J.-E. Muller, U. Ablassmeier, J. Schelle, W. Kellner, and H. Kniepkamp, "A double-conversion broad band TV-tuner with GaAs ICs," in *GaAs IC Symp.*, 1984, pp. 97–100.
[167] P. Dautriche, B. Y. Lao, C. Villalon, V. Pauker, M. Bostelmann, and M. Binet, "GaAs monolithic circuits for TV tuners," in *GaAs IC Symp.*, Monterey, CA, 1985, pp. 165–168.
[168] J. Fenk and R. Tauber, "TV VHF/hyperband tuner ICs," *IEEE Trans. Consum. Electron.*, vol. CE-32, pp. 723–733, Apr. 1986.
[169] H. van Glabbeek, N. Baars, H. Schreurs, and W. Zwijsen, "VHF, hyperband, and UHF mixer/oscillator sections of a TV tuner on one IC," *IEEE Trans. Consum. Electron.*, vol. CE-33, pp. 619–622, Apr. 1987.
[170] T. Nakatsuka, S. Sakashita, K. Goda, and S. Nambu, "A GaAs RF amplifier IC for UHF TV tuners," *IEEE Trans. Consum. Electron.*, vol. 34, pp. 366–371, Feb. 1988.
[171] H. Mizukami, K. Sakuta, H. Hatashita, T. Nagashima, and K. Shinkawa, "A high quality GaAs IC tuner for TV/VCR receivers," *IEEE Trans. Consum. Electron.*, vol. 34, pp. 649–658, Mar. 1988.
[172] T. Ducourant, P. Philippe, P. Dautriche, V. Pauker, C. Villalon, M. Pertus, and J.-P. Damour, "A 3-chip GaAs double-conversion TV tuner with 70 dB image rejection," in *Microwave & Millimeter-Wave Monolithic Circ. Symp.*, Long Beach, CA, 1989, pp. 87–90.
[173] H. Yagita, A. Terao, M. Tsuneoka, A. Watanabe, T. Tambo, and S. Nambu, "Low noise and low distortion GaAs mixer-oscillator IC for broadcasting satellite TV tuner," in *GaAs IC Symp.*, San Diego, CA, 1989, pp. 75–78.
[174] P. Philippe and M. Pertus, "A 2 GHz enhancement-mode GaAs down-converter IC for satellite TV tuner," in *Microwave Theory & Techniques Symp.*, Boston, MA, 1991, pp. 73–76.
[175] H. Mizukami, H. Ikedo, K. Ideno, T. Nagashima, and S. Yamada, "Low supply voltage GaAs ICs for a TV tuner," in

Int. Symp. on GaAs and Related Compounds, Karuizawa, Japan, 1992.
[176] K. Ideno, H. Mizukami, T. Nagashima, and K. Sakamoto, "A wide band GaAs IC for a DBS/TV receiver," in *Int. Symp. on GaAs and Related Compounds*, Karuizawa, Japan, 1992.
[177] K. Washio, T. Okabe, K. Norisue, and T. Nagashima, "An all-band TV tuner IC with 10 GHz 100 V mixed analog/digital Si bipolar technology," *IEEE J. Solid-State Circ.*, vol. 27, pp. 1264–1269, Sept. 1992.
[178] N. Scheinberg, R. Michels, V. Fedoroff, and D. Stoffman, "A GaAs upconverter integrated circuit for a double conversion cable TV set-top tuner," *IEEE J. Solid-State Circ.*, vol. 29, pp. 688–692, June 1994.
[179] J. Hurn, *GPS: A Guide to the Next Utility*. Sunnyvale, CA: Trimble Navigation, 1989.
[180] W. O. Dussell and J. Medina, "Using creative surface mount manufacturing and packaging as a competitive weapon," in *Surface Mount Int. Conf. & Expo.*, San Jose, CA, 1993, pp. 241–245.
[181] R. M. Herman, C. H. Mason, H. P. Warren, and R. A. Meier, "A GPS receiver with synthesized local oscillator," in *Int. Solid-State Circ. Conf.*, New York, 1989, pp. 194–195.
[182] R. M. Herman, A. Chao, C. H. Mason, and J. R. Pulver, "An integrated GPS receiver with synthesizer and downconversion functions," in *Int. Microwave Symp.*, Boston, 1991, pp. 883–886.
[183] R. Benton, M. Nijjar, C. Woo, A. Podell, G. Horvath, E. Wilson, and S. Mitchell, "GaAs MMICs for an integrated GPS front-end," in *GaAs IC Symp.*, Miami Beach, FL, 1992, pp. 123–126.
[184] K. J. Negus, R. A. Koupal, D. Millicker, and C. P. Snapp, "3.3 V GPS receiver MMIC implemented on a mixed-signal, silicon bipolar array," in *Microwave & Millimeter-Wave Monolithic Circ. Symp.*, Albuquerque, NM, 1992, pp. 209–212.
[185] N. Camilleri, J. Kirschgessner, J. Costa, D. Ngo, and D. Lovelace, "Bonding pad models for silicon VLSI technologies and their effects on the noise figure of RF NPNs," in *Microwave & Millimeter-Wave Monolithic Circ. Symp.*, San Diego, CA, 1994, pp. 225–228.

Asad A. Abidi (Member, IEEE) was born in 1956. He received the B.Sc. (Hon.) degree from Imperial College, London in 1976 and the M.S. and Ph.D. degrees in electrical engineering from the University of California, Berkeley in 1978 and 1981.

From 1981 to 1984, he was at Bell Laboratories, Murray Hill, NJ, as a Member of Technical Staff in the Advanced LSI Development Laboratory. Since 1985, he has been at the Electrical Engineering Department of the University of California, Los Angeles, where he is a Professor. He was a Visiting Faculty Researcher at Hewlett Packard Laboratories during 1989. His research interests are in CMOS RF design, high-speed analog integrated circuit design, data conversion, and other analog signal processing techniques.

Dr. Abidi served as the Program Secretary for the International Solid State Circuits Conference from 1984 to 1990, and as General Chairman of the Symposium on VLSI Circuits in 1992. He was Secretary of the IEEE Solid State Circuits Council from 1990 to 1991, and is now Editor of the *IEEE Journal of Solid State Circuits*. He received the 1988 TRW Award for Innovative Teaching.

EM Interaction of Handset Antennas and a Human in Personal Communications

M.A. Jensen and Y. Rahmat-Samii

Proceedings of the IEEE
vol. 83, no. 1, pp. 7-17
January 1995

EM Interaction of Handset Antennas and a Human in Personal Communications

MICHAEL A. JENSEN, STUDENT MEMBER, IEEE, AND YAHYA RAHMAT-SAMII, FELLOW, IEEE

In personal communications, the electromagnetic interaction between handset-mounted antennas and the nearby biological tissue is a key consideration. This paper presents a thorough investigation of this antenna-tissue interaction using the finite-difference time-domain (FDTD) electromagnetic simulation approach with detailed models of real-life antennas on a transceiver handset. The monopole, side-mounted planar inverted F, top-mounted bent inverted F, and back-mounted planar inverted F antennas are selected as representative examples of external and internal configurations. Detailed models of the human head and hand are implemented to investigate the effects of the tissue location and physical model on the antenna performance. Experimental results are provided which support the computationally obtained conclusions. The specific absorption rate (SAR) in the tissue is examined for several different antenna/handset configurations. It is found that for a head-handset separation of 2 cm, the SAR in the head has a peak value between 0.9 and 3.8 mW/g and an average value between 0.06 and 0.10 mW/g for 1 W of power delivered to the antenna. Additionally, the head and hand absorb between 48 and 68% of the power delivered to the antenna.

I. INTRODUCTION

The introduction of cellular communications technology has generated a widespread awareness of the important role wireless services play in today's communications-centered marketplace [1]. Naturally, this awareness has spawned a growth in the wireless communications arena which in turn has had and will continue to have an impact on antenna technology advancement. An important example of this influence involves the development of external and internal antenna structures, such as the monopole or planar inverted *F* antenna (PIFA), which can be efficiently and conveniently integrated with a hand-held transceiver unit. The dominant role of such portable terminals in many systems motivates a thorough examination of the issues involved in designing these and other similar antenna structures.

Because portable handsets operate in close proximity to a human being, one particularly important consideration involves the interaction of the radiated electromagnetic fields with the nearby biological tissue. Certainly, the operator's influence on the antenna gain, radiation pattern, and input impedance is an issue which deserves a detailed investigation. Furthermore, growing concerns over the health effects of tissue exposed to electromagnetic energy motivates an effort to understand the power absorption distribution in the tissue when a hand-held device is used.

To gain a detailed understanding of the antenna-tissue interaction, the designer must abandon simplified, approximate analysis methodologies in favor of more general numerical simulation techniques. The finite-difference time-domain (FDTD) method [2]–[6] is one such approach which allows accurate prediction of the electromagnetic behavior of general topologies. A few studies have appeared which present preliminary results associated with this topic. For example, one paper has used a homogeneous sphere and block of muscle to represent the head and hand, respectively, within an FDTD simulation of a handset-mounted monopole antenna [7]. Other studies have used a more elaborate model of the head, but have used simpler models for the antenna such as a dipole [8].

In this work, advanced physical models of the head and hand are coupled with detailed representations of handset-mounted antennas to allow accurate FDTD simulations of real-life communications scenarios. The computer program developed for this work has been made adequately flexible to accommodate these modeling requirements. The computations not only reveal the human operator influence on the antenna performance, but also provide information concerning the SAR [8]–[10] for power absorbed in the tissues. The monopole, side-mounted PIFA, top-mounted bent inverted *F* antenna (BIFA), and back-mounted PIFA elements shown in Fig. 1 are used in the computational examples to provide results representative of external and internal antennas. The effect of tissue location and the model used to represent the human operator are addressed. Measured data is compared to many of the computational results to show the accuracy of the FDTD tool in predicting the antenna performance. Conclusions and suggestions are presented based upon the results obtained.

II. PHYSICAL/COMPUTER MODELING

A. FDTD Implementation

The FDTD methodology [2]–[10] is derived from Maxwell's time-domain equations which may be expressed

Manuscript received February 22, 1994; revised August 30, 1994. This work was supported by ARPA Contract Number DAAB07-93-C-C501.

The authors are with the Electrical Engineering Department, University of California at Los Angeles, Los Angeles, CA 90024-1594 USA.

IEEE Log Number 9406485.

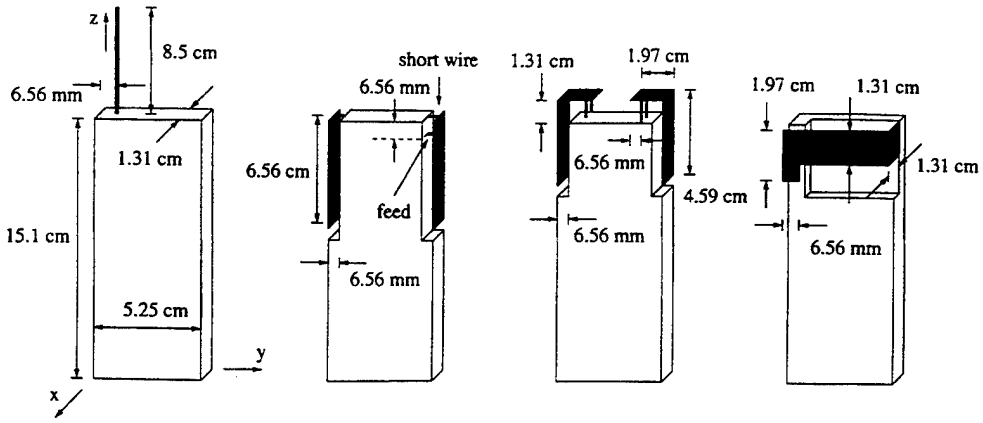


Fig. 1. Antenna geometries and dimensions for hand-held transceiver units: (a) monopole; (b) side-mounted dual PIFA; (c) top-mounted dual BIFA; (d) back-mounted PIFA. The chassis dimensions (105 cm^3) shown in (a) apply to all four configurations.

as

$$\mu\frac{\partial\vec{H}}{\partial t} = -\nabla\times\vec{E} \tag{1}$$

$$\epsilon\frac{\partial\vec{E}}{\partial t} = \nabla\times\vec{H} - \sigma\vec{E} \tag{2}$$

where $\vec{E}$ and $\vec{H}$ are the electric and magnetic field intensities and ϵ, μ, and σ are the space-dependent permittivity, permeability, and conductivity, respectively. Using a discretization of the calculus operators in these equations results in a set of algebraic *time-stepping* relations which may be written in a compact form as

$$\vec{H}^{n+1/2} = \vec{H}^{n-1/2} + \overline{\overline{\gamma}}\cdot\left\{D_x\left(-\hat{x}\times\vec{E}^n\right) + D_y\left(-\hat{y}\times\vec{E}^n\right) + D_z\left(-\hat{z}\times\vec{E}^n\right)\right\} \tag{3}$$

$$\vec{E}^{n+1} = \overline{\overline{\alpha}}\cdot\vec{E}^n + \overline{\overline{\beta}}\cdot\left\{D_x\left(\hat{x}\times\vec{H}^{n+1/2}\right) + D_y\left(\hat{y}\times\vec{H}^{n+1/2}\right) + D_z\left(\hat{z}\times\vec{H}^{n+1/2}\right)\right\} \tag{4}$$

where the superscript denotes the time step $t = n\Delta t$ and the difference operator is defined as

$$D_x\,\vec{f}(x,y,z) = \frac{\vec{f}(x+\Delta x/2,y,z) - \vec{f}(x-\Delta x/2,y,z)}{\Delta x}. \tag{5}$$

The terms $\overline{\overline{\alpha}}$, $\overline{\overline{\beta}}$, and $\overline{\overline{\gamma}}$ are space-dependent diagonal tensors whose components are defined by

$$\alpha_{pp} = \frac{\frac{\bar{\epsilon}_p}{\Delta t} - \frac{\bar{\sigma}_p}{2}}{\frac{\bar{\epsilon}_p}{\Delta t} + \frac{\bar{\sigma}_p}{2}} \qquad \beta_{pp} = \frac{1}{\frac{\bar{\epsilon}_p}{\Delta t} + \frac{\bar{\sigma}_p}{2}} \qquad \gamma_{pp} = \frac{\Delta t}{\bar{\mu}_p} \tag{6}$$

for $p = x, y$, or z. The symbols characterized by an overbar ($\bar{\cdot}$) represent averaged values of the constitutive parameters over a face of a cell in the discretized computational grid, such as

$$\bar{\epsilon}_p = \frac{\int_{\Delta S}\epsilon(\vec{r})\hat{p}\cdot d\vec{S}}{\Delta S} \tag{7}$$

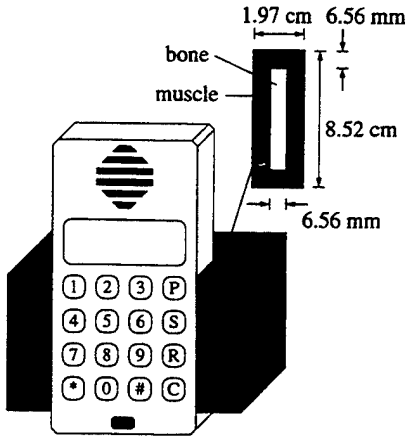


Fig. 2. Block model of the human hand used in the simulation of antennas on the handset.

where ΔS represents the area of the cell face normal to the unit vector $\hat{p}$. Using the definitions in (6) and (7) allows accurate modeling of inhomogeneous materials within the computational domain.

The relations in (3) and (4) are used to track the time-evolution of the fields in the spatial domain by initially setting all field values to zero. An antenna excitation is then introduced by specifying a voltage at the antenna feed point. The value of the magnetic and electric field intensities are alternately computed at times $t = (n+1/2)\Delta t$ and $t = (n+1)\Delta t$. At grid points coincident with perfect electric conductors, the tangential components of $\vec{E}$ are set to zero at each time step. Special considerations are made for modeling wires and lumped elements within the domain [5], and absorbing boundary conditions [11] are used to truncate the mesh a reasonable distance from the antenna. Once the time-domain data has been collected, a Fourier transform is used to obtain the desired frequency-domain quantities such as input impedance, radiation pattern, and gain. In this fashion, the antenna response over a given bandwidth may be obtained.

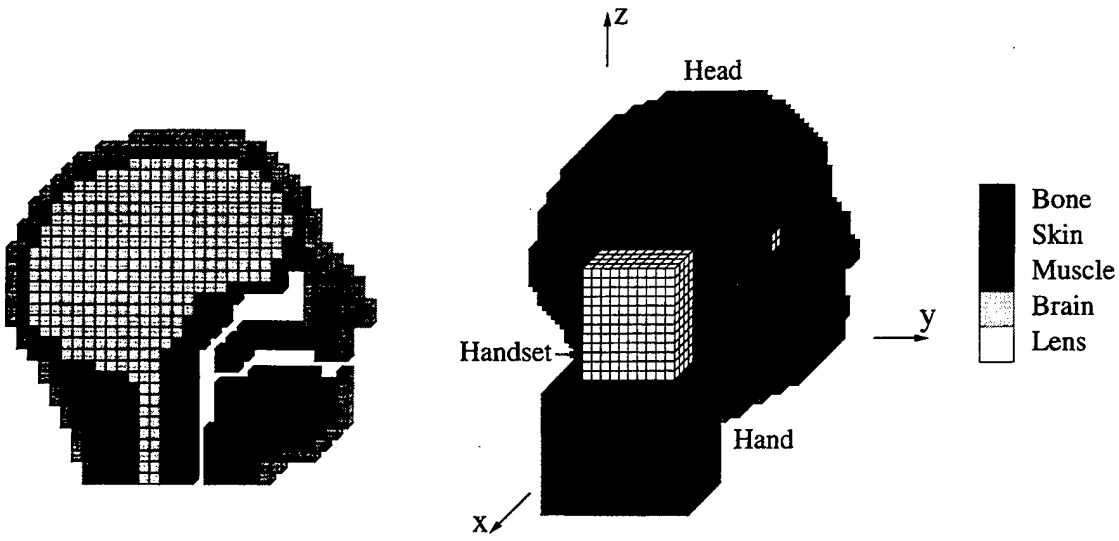


Fig. 3. (a) Sagittal cut of the discrete head model at the head center and (b) entire head model with the hand and plastic-covered handset present.

The antenna excitation is introduced into the computational domain using a previously reported simulated coaxial feed model [5]. The functional form of the voltage applied in the feed is a sinusoid modulated by a Gaussian as expressed by

$$s(t) = e^{-(t-t_o)^2/2\sigma_t^2} \cos[2\pi f_o(t-t_o)] \qquad (8)$$

where f_o is the sinusoidal frequency and σ_t controls the pulse width. This particular functional form is interesting since the center frequency and bandwidth of its spectrum may be controlled. In this work, antenna impedance computations over a band are performed using the parameters σ_t= 160 ps and f_o= 0. In computing this transient response, the time-stepping is continued until the antenna current magnitude has decayed to a level approximately 40 dB below its peak value. When computing radiation patterns and power absorption in tissues, a single frequency source is used such that $\sigma_t \rightarrow \infty$. In this case, the time-stepping is performed for several cycles (~20) in order for the system to reach a steady state.

B. Biological Tissue Modeling

Anatomical human features are modeled within the FDTD framework by mapping the spatial location of the different tissues into a permittivity and conductivity assignment in the computational grid. The human hand is simply modeled as a layer of bone surrounded by a layer of muscle that covers three sides of the handset, as depicted in Fig. 2. To construct a head model, a grid with a 6.56 mm spatial resolution was placed on cross-sectional images of the head obtained from an anatomy atlas [12]. Magnetic resonance images (MRI) were also used to aid in the tissue classification and location. Each cell in the grid was then assigned a permittivity and conductivity classification corresponding to the type of tissue which filled the *majority* of the cell. Figs. 3(a) and (b) illustrate a mid-sagittal cross section of the head model and the full head model (24 × 33 × 31 cells) with the hand and handset. Fig. 4 provides two different views of the head/hand/handset configuration with the dimensions used in the computations. The electrical parameters corresponding to each tissue around the operating frequency of 915 MHz have been obtained from published data and are listed in Table 1 [9]. Since variations in the conductivity and permittivity values are less than 5% and 1%, respectively, over the frequency band of interest, these constitutive parameters will be considered to be nondispersive in the FDTD simulations. If broadband computations are to be performed where the dispersive nature of the tissues is more significant, then special considerations in the FDTD implementation must be made [13].

Two additional comments are warranted concerning the tissue models presented. First, in order to allow the modeling of very realistic operator/handset configurations, it is important to allow the handset to be rotated such that it is positioned between the operator's mouth and ear. This is accomplished by rotating the position of each cell in the head model in the $y-z$ plane about the head center and reconstructing the grid based upon these rotated tissue locations (see Fig. 4). The fact that the handset remains aligned with the FDTD grid allows accurate modeling of its rectilinear features. Second, when using the models in the computations, different FDTD cell sizes are chosen according to the parameter of interest. For example, it has been found that a cell size of 3.28 mm is required in order to obtain an accurate value of the input impedance. However, for pattern and power absorption characterization, a 6.56 mm cell size may be used. These two cell sizes correspond to approximately 14 and 7 cells per wavelength,

Table 1. Relative Permittivity, Conductivity, and Density of the Tissues in the Hand and Head Near 900 MHz

Tissue	Permittivity	Conductivity (S/m)	Density (g/cm^3)
Bone	8.0	0.105	1.85
Skin/Fat	34.5	0.60	1.10
Muscle	58.5	1.21	1.04
Brain	55.0	1.23	1.03
Humour	73.0	1.97	1.01
Lens	44.5	0.80	1.05
Cornea	52.0	1.85	1.02

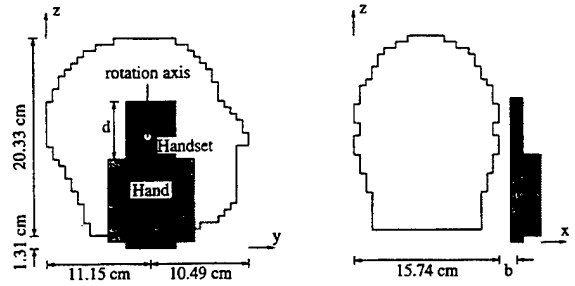


Fig. 4. Side and rear views of the FDTD head/hand/handset model showing dimensions.

respectively, *within* the high-permittivity tissues at the operating frequency.

C. Time-Response of High-Permittivity Dielectrics

The relatively high permittivity values listed in Table 1 suggest that a considerable amount of computational time (i.e., large number of time steps) is required for an antenna near a biological system to reach its steady-state response for a given input. To test the duration of the transient effects, a 16.4 cm dipole is placed 1.31 cm from a 17 cm diameter homogeneous dielectric sphere. The antenna is excited with a 160 ps Gaussian pulse and the antenna terminal current I_s is monitored as a function of time for different values of the sphere permittivity and conductivity. Fig. 5 illustrates the result of this study, where the normalization parameters η_o and c_o are the free-space impedance and speed of light, respectively. The dielectric parameters are chosen to represent the humor in the eye since this is the highest-permittivity tissue involved in the models. As can be seen, introduction of high-permittivity dielectrics produces significant ringing in the antenna response. However, addition of loss into the system causes the transient response to damp quickly. The response for the dipole near the head model also shown in Fig. 5 illustrates that this same damping effect occurs for the inhomogeneous tissue models used in this investigation.

D. Power Absorption and SAR

Several quantities of interest should be defined before proceeding with the results. The power absorbed within the

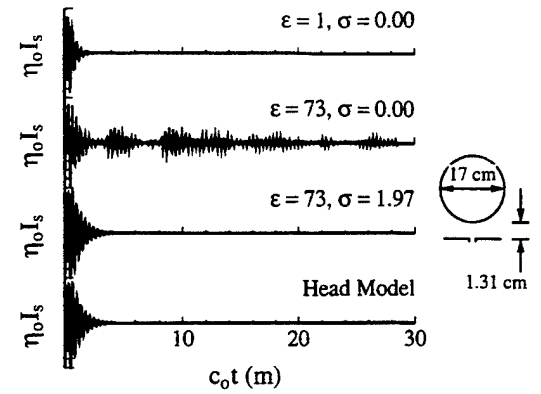


Fig. 5. Normalized antenna current versus normalized time for a dipole 1.31 cm from a homogeneous sphere filled with dielectrics of different permittivity and conductivity values. The case for a dipole next to the inhomogeneous head model is also provided.

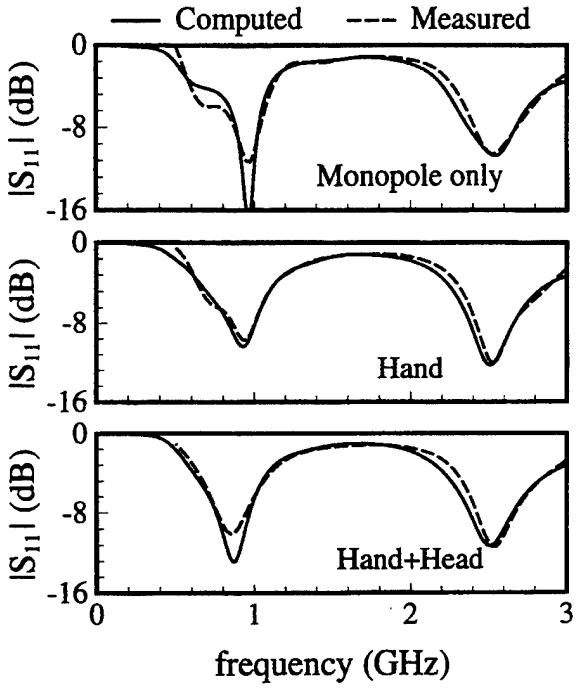


Fig. 6. Computed and measured values of $|S_{11}|$ for the monopole on the handset with no tissue, with the hand only, and with the hand and head (b = 1.97 cm, d = 5.9 cm).

lossy tissue (P_{abs}) and the power radiated to the far-field (P_{rad}) are obtained using the integrals

$$P_{\text{abs}} = \frac{1}{2}\int_V \sigma |\mathbf{E}|^2 dV \tag{9}$$

$$P_{\text{rad}} = \frac{1}{2}\,\text{Re}\left\{\int_S \mathbf{E}\times\mathbf{H}^*\cdot\hat{n}\,dS\right\} \tag{10}$$

where $\mathbf{E}$ and $\mathbf{H}$ are the frequency domain peak electric and magnetic fields, respectively, and σ is the medium conductivity. The symbols V, S, and $\hat{n}$, represent the volume containing the tissue, the surface completely surrounding the handset/operator configuration, and the unit outward normal to the surface, respectively. The antenna/tissue

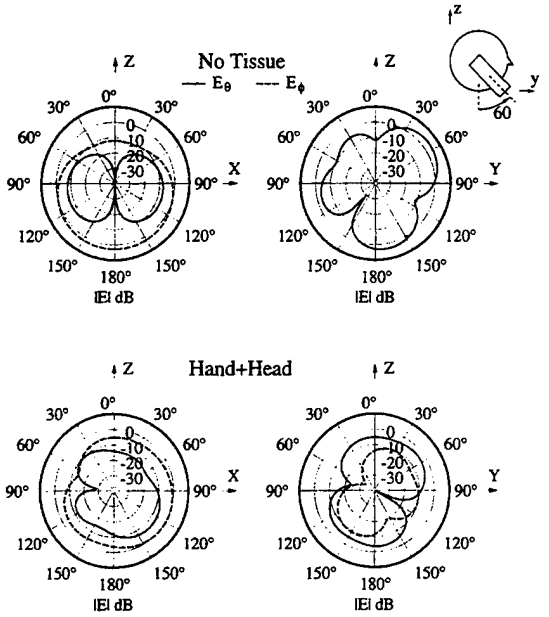


Fig. 7. Computed radiation patterns at 915 MHz normalized to the antenna gain for a monopole on the plastic-covered handset in the absence and presence of the head and hand (b = 1.97 cm, d = 5.9 cm). The antennas are rotated 60° from upright.

system radiates with an efficiency of

$$\eta_a = \frac{P_{\text{rad}}}{P_{\text{rad}} + P_{\text{abs}}} = \frac{P_{\text{rad}}}{P_{\text{del}}} \tag{11}$$

where P_{del} is the real power delivered to the antenna.

The SAR quantifies the power absorbed per unit mass of tissue and is a fundamental parameter when discussing the health risks of electromagnetic power absorption in the body. This quantity is defined as

$$\text{SAR} = \frac{\sigma}{2\rho}|\mathbf{E}|^2 \tag{12}$$

where ρ is the material density. The values of σ and ρ for the different tissue types are provided in Table 1 [9]. The ANSI/IEEE standard suggests that the SAR averaged over any 1 g of tissue for 30 min. or more should remain under 1.6 mW/g [14], [15]. If necessary, the SAR values can be manipulated to obtain the tissue heating resulting from the power dissipation through a simple multiplicative constant.

III. Computational and Experimental Results

The following cases provide examples of computations and measurements used in the evaluation of several antenna/handset configurations. The experimental measurements of S_{11} provided here have been obtained from a Hewlett–Packard 8510B network analyzer at the University of California, Los Angeles, antenna measurement facility. This parameter is a measure of the reflection coefficient for a voltage wave introduced to the antenna feed point using a 50 Ω coaxial cable and may be defined as

$$S_{11} = \frac{Z_a - 50}{Z_a + 50} \tag{13}$$

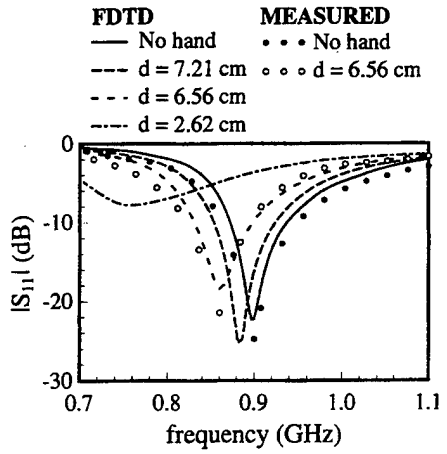


Fig. 8. Computed value of $|S_{11}|$ for the side-mounted PIFA on the handset without the hand and with the hand for three different values of d (see Fig. 4). Measured values appear for the configurations with no hand and with the hand at d = 6.56 cm.

where Z_a is the antenna input impedance. For all of the examples, the feeding wire radius is r_o = 0.46 mm which corresponds to the inner conductor radius of RG402/U 50 Ω semi-rigid coaxial cable. Where impedance measurements involving biological tissue are presented, a single human subject is used. These experiments have been deliberately performed at very low power levels such that induced SAR levels are well below recommended safety limits.

The computer platform used in the simulations is an IBM RISC/6000 530H workstation. Based upon the findings of Section II-C, the simulations for the small (large cell size) and large (small cell size) models require 1800 and 3600 time steps, respectively, corresponding to respective physical run times of approximately 2–3 hours and 10–12 hours. The storage requirements are approximately 17 MB for the small model and 35 MB for the large model.

The geometries and dimensions for the different antennas under investigation are illustrated in Fig. 1. Parts (a)–(d) illustrate the monopole, side-mounted PIFA, top-mounted BIFA, and back-mounted PIFA configurations, respectively. The chassis dimensions provided in Fig. 1(a) apply to all four geometries. When the plastic casing is included, it is modeled as a 3.28 mm thick lossless dielectric with a relative permittivity of 2 which is immediately adjacent to and completely surrounds the chassis. In computations where the FDTD cell size is larger than 3.28 mm, the approach detailed in [16] is used to approximate the thin plastic casing.

A. Monopole Antenna

The monopole antenna's widespread use in the personal wireless industry motivates an investigation of its characteristics and performance. The geometry for the quarter-wave monopole antenna on the conducting handset chassis is shown in Fig. 1(a). Fig. 6 presents the value of $|S_{11}|$ versus frequency for this configuration obtained both computationally and experimentally for the handset alone, the handset with the hand (d = 5.9 cm), and the

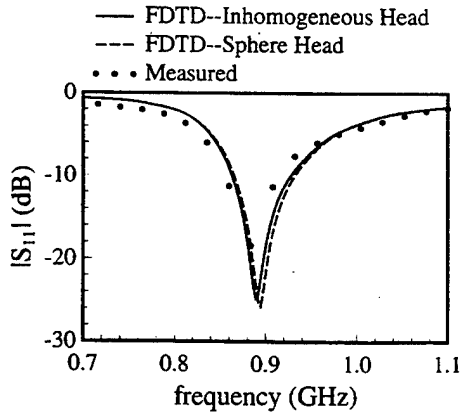


Fig. 9. $|S_{11}|$versus frequency for the side-mounted PIFA on the handset with the head and the hand (b = 1.97 cm, d = 7.21 cm). Computed results obtained using both head models are compared with measured data.

handset with the hand and the head (b = 1.97 cm). For simplicity, the data for the third case is obtained with the handset oriented upright with respect to the head, as illustrated in Fig. 4. Because no plastic casing was included in the experimental handset prototype, the casing was absent in the computations as well. As can be seen, the FDTD method is very effective in predicting the antenna impedance behavior. Furthermore, the results imply that the hand and head exercise relatively little influence on the antenna input impedance. This is an important feature of the monopole which partly accounts for its widespread use.

The monopole's radiation pattern, in contrast to its input impedance, is more noticeably influenced by the operator proximity. This is demonstrated in Fig. 7 which provides the radiation pattern of the handset-mounted monopole at 915 MHz in the absence and presence of the hand and head. In this simulation, the plastic-covered handset is rotated 60° in the $y-z$ plane with respect to the head (see figure inset), and the patterns are normalized to the antenna gain. These patterns are presented in standard spherical coordinates such that the principal polarizations are in the $\hat{\theta}$ and $\hat{\phi}$ directions. Clearly, the presence of the biological tissue alters the radiation pattern and reduces the antenna gain. The ratio of the power absorbed in the head and hand to the total power delivered to the antenna, as well as the antenna radiation efficiency, are provided in Table 2 for the 915 MHz operation frequency. This table contains data for the handset rotated as in Fig. 7 and for the handset upright as shown in Fig. 4. As can be seen, the results are very similar for the two handset orientations. The gain reduction caused by power absorption is a key issue to consider in determining system requirements to satisfy target link budgets.

B. Side-Mounted PIFA

The growing desire to replace the monopole with more conformal, less obtrusive elements has focussed considerable attention on antennas such as the PIFA [17]. A potential topology with two PIFA elements configured

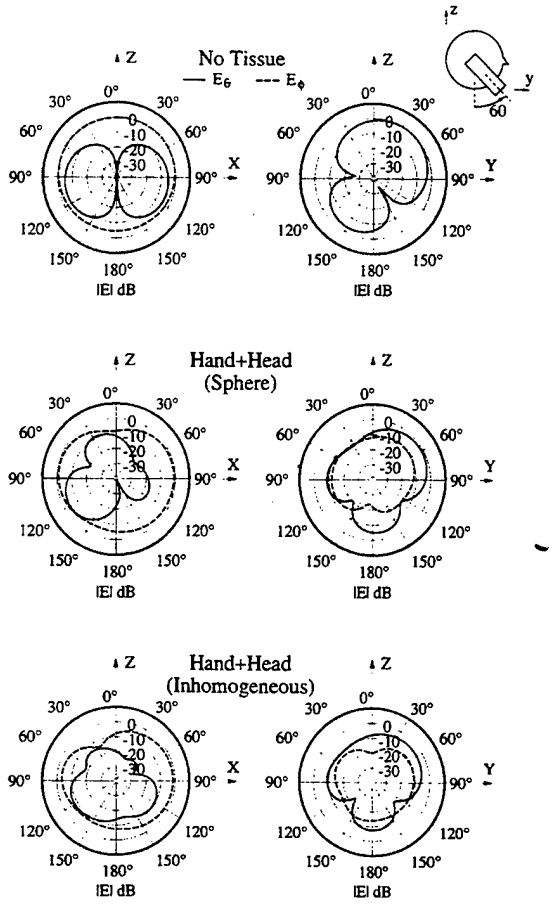


Fig. 10. Computed patterns at 915 MHz normalized to the antenna gain for the side-mounted PIFA on the plastic-covered handset rotated 60° from upright. Results are shown for the handset alone, with the spherical head model, and with the inhomogeneous head model (b = 1.97 cm, d = 7.21 cm).

to provide dual-branch antenna diversity is illustrated in Fig. 1(b) [5]. These side-mounted elements consist simply of a probe-fed conducting plate suspended above the conducting chassis. A short circuiting wire or strip is attached to one end of the suspended plate—a configuration which allows considerable reduction in the element resonant size. The antenna can generally be matched to a 50 Ω feeding line through proper selection of the feed point, shorting pin location, and antenna dimensions.

Fig. 8 illustrates the value of $|S_{11}|$ versus frequency for one PIFA element with the second terminated in a matched load for several handset/operator configurations. The curves compare the results when no tissue is included to those when the hand is placed at three vertically displaced positions on the handset. Once again, no plastic casing is included in order to allow comparison with measured results. As can be seen from the plot, the hand position exercises a detuning effect on the antenna resonant frequency and impedance. Most significant is the high impedance mismatch which occurs when the hand begins to mask the antenna. These results illustrate the importance of minimizing antenna masking through proper antenna placement. The dots in the figure indicate experimentally

Table 2. Computed Normalized Power Absorption and Peak SAR (mW/g/W) in the Head and Hand, Average SAR (mW/g/W) in the Head, and Radiation Efficiency for the Different Handset/Body Configurations Shown in Fig. 13. All Data is Computed at 915 MHz

Configuratio	$\frac{P_{abs}^{head}}{P_{del}}$	$\frac{P_{abs}^{hand}}{P_{del}}$	η_a	$\frac{SAR_{max}^{head}}{P_{del}}$	$\frac{SAR_{max}^{hand}}{P_{del}}$	$\frac{SAR_{ave}^{head}}{P_{del}}$
Handset Rotated 60° wuth Respect to Head						
Monopole	0.359	0.169	0.472	1.97	2.29	0.0875
Side Mounted PIFA	0.382	0.301	0.317	3.81	4.54	0.0931
Side Mounted PIFA†	0.212	0.318	0.470	3.14	5.42	0.1003
Back Mounted PIFA	0.260	0.222	0.518	1.32	3.58	0.0634
Handset Upright with Respect to Head						
Monopole	0.351	0.184	0.465	2.06	2.43	0.0856
Side Mounted PIFA	0.332	0.324	0.344	2.07	4.91	0.0809
Back Mounted PIFA	0.258	0.225	0.517	0.90	3.53	0.0629

†Homogeneous spherical head (ϵ_r=58.5, σ=1.21)

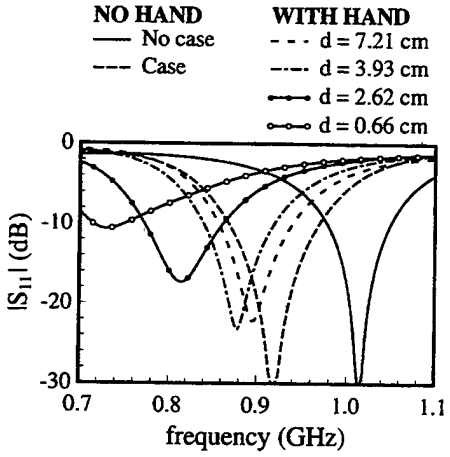


Fig. 11. Computed value of $|S_{11}|$ for the top-mounted BIFA on the handset without the plastic casing, with the casing, and with the casing and hand at four different locations.

measured data for the antenna with the hand absent and present at d = 6.56 cm. This comparison shows good correlation between the experimentally and computationally obtained results.

The curves in Fig. 9 represent $|S_{11}|$ when the head and hand are present for two different head models. The curve labeled 'sphere' corresponds to the case of a homogeneous spherical ball of muscle (ϵ_r =58.5, σ = 1.21 S/m) with a radius of 9 cm. The second head is the inhomogeneous model discussed in Section II-B. As can be seen, for this particular configuration the choice of models exercises little influence on the antenna input impedance. This insensitivity occurs because the input impedance is a reasonably local phenomenon which is influenced most significantly by structures in the near vicinity of the antenna rather than objects such as the head which are displaced somewhat from the antenna feed point. Once again, the measured data given in Fig. 9 shows that the FDTD accurately predicts the effects of the human operator on the antenna performance.

Unlike the input impedance, the radiation characteristics of the side-mounted PIFA are expected to show an increased sensitivity to the presence of the head as well as the physical model used. This expectation is confirmed in Fig. 10 which shows the radiation patterns normalized to the antenna gain at 915 MHz when the handset is encased in plastic and rotated 60° with respect to the head. The three sets of curves in Fig. 10 represent the patterns for the handset alone, the handset near the spherical head, and the handset near the detailed head (b = 1.97 cm). When either head model is used, the hand is placed at d = 7.21 cm. As anticipated, the antenna radiation pattern depends on the tissue presence as well as the choice of head models. Once again, the amount of power absorbed in the tissue is quantified in Table 2. As can be seen, the inhomogeneous head model results in somewhat higher absorption levels as compared to the simpler spherical head model for this configuration. This difference occurs because the sphere is physically smaller and has a permittivity/conductivity distribution which reduces the electromagnetic penetration depth as compared to the inhomogeneous model.

C. Top-Mounted BIFA

As mentioned above, proper antenna placement is a critically important issue when designing antennas for hand-held devices. The goal is to find an antenna location to minimize radiation blockage by the hand. In general, observation reveals that the typical user will not hold the transceiver near the top. For this reason, the configuration shown in Fig. 1(c) is investigated. The variation of $|S_{11}|$ with frequency is shown in Fig. 11. The different curves

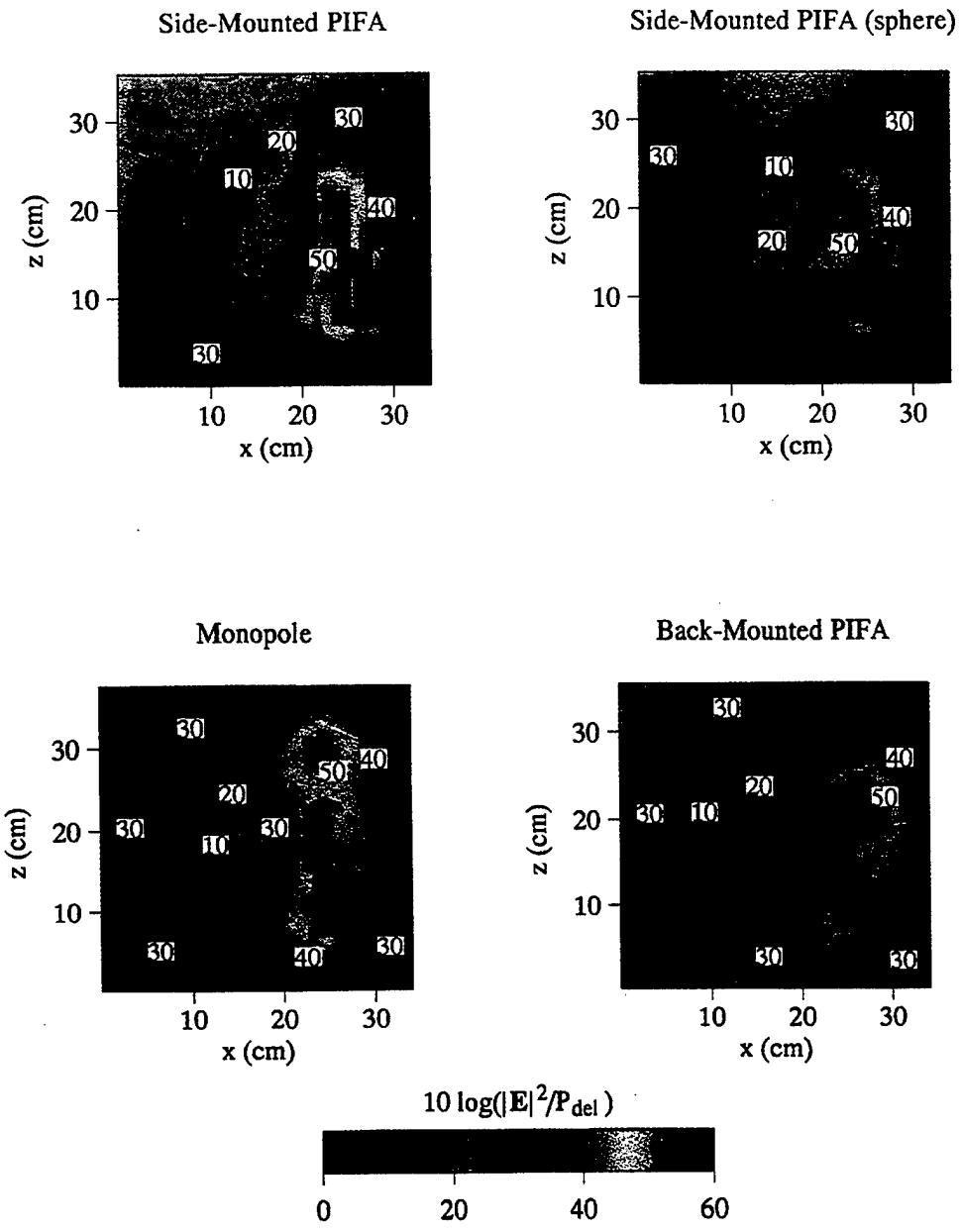


Fig. 12. Computed normalized near field distribution in a plane through the center of the head with d = 7.21 cm, b = 1.97 cm, and the plastic casing present at 915 MHz. The configurations are the side-mounted PIFA, the side-mounted PIFA with the spherical head, the monopole, and the back-mounted PIFA.

represent computations for the handset alone, with the casing, and with the casing and the hand at four vertically displaced positions. Consistent with the results for the side-mounted PIFA, the presence of the hand detunes the antenna resonance. However, in comparison with the geometry of Fig. 1(b), the degradation for the top-mounted BIFA is less severe for a given hand position. For this reason, configurations similar to this show potential as possible integrated antennas for hand-held transceiver units. The results in Fig. 11 also reveal the significant frequency detuning caused by the plastic casing. Compensation for this effect should be included in the integrated antenna design.

D. Near Fields and Specific Absorption Rate

Having investigated the effect of the body on the electromagnetic characteristics of these antenna structures, we now turn our focus to the field distribution and power absorption characteristics within the tissue. Fig. 12 compares the field variation around the plastic-covered handset, the head, and the hand at 915 MHz for the side-mounted PIFA with the inhomogeneous and spherical head models and for the monopole antenna. Fig. 13 presents the SAR distribution for the same configurations. For each computation, d = 7.21 cm, b = 1.97 cm, and the head is upright for simplicity in data presentation. The data plane is located at the center

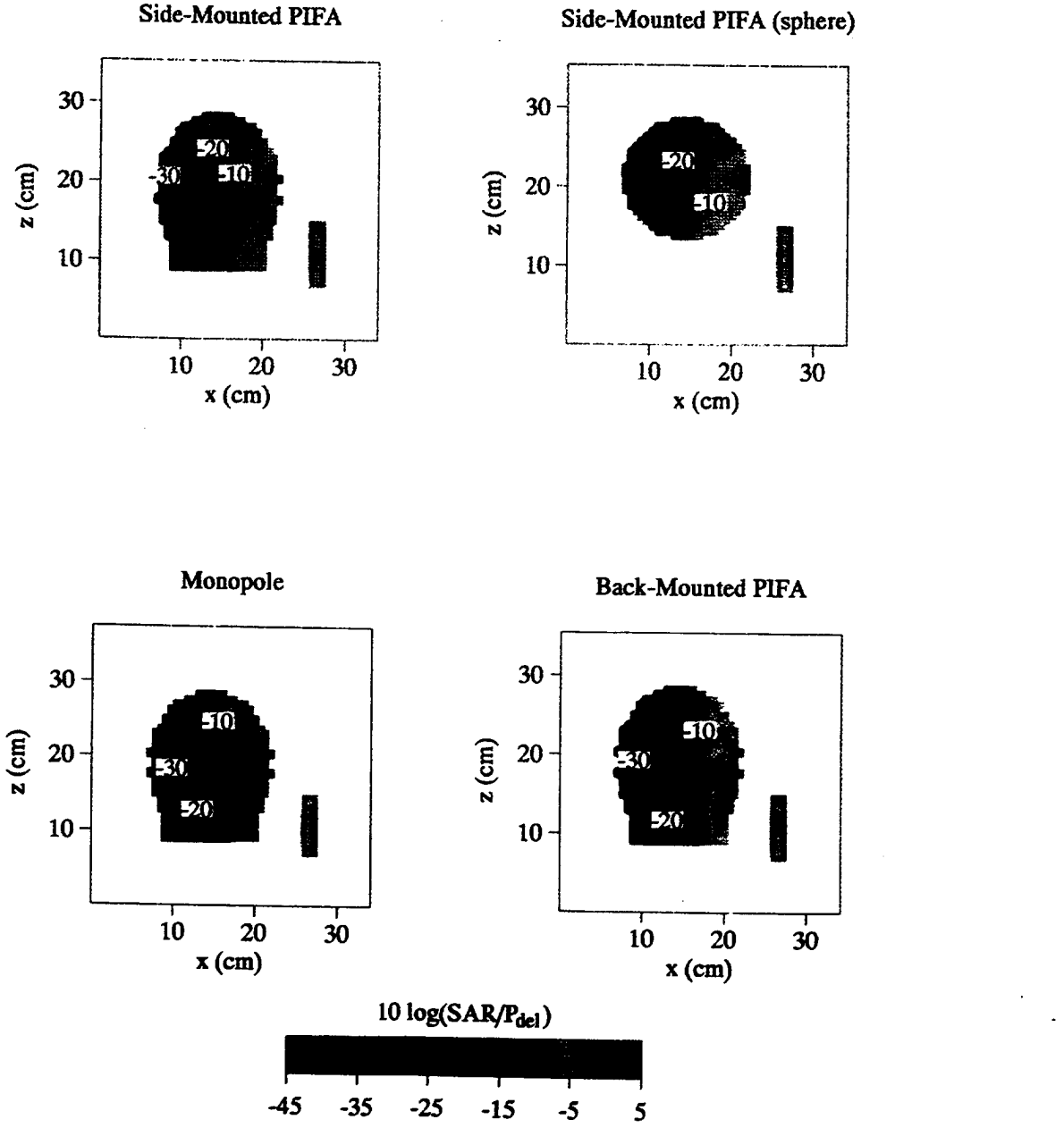


Fig. 13. Computed normalized SAR distribution in a plane through the center of the head with d = 7.21 cm, b = 1.97 cm, and the plastic casing present at 915 MHz. The configurations are the side-mounted PIFA, the side-mounted PIFA with the spherical head, the monopole, and the back-mounted PIFA.

of the head/handset combination, and the plots are viewed from the $-y$ direction (see Fig. 4). The values in Fig. 12 represent the total squared magnitude of the peak electric field per watt of total power delivered to the antenna, or

$$10 \log \frac{|\mathbf{E}|^2}{P_{\text{del}}}. \tag{14}$$

Similarly, the values in Fig. 13 represent the SAR averaged over 1 g of tissue and normalized to the power delivered to the antenna using the expression

$$10 \log \frac{\text{SAR}}{P_{\text{del}}}. \tag{15}$$

Table 2 provides a numerical comparison of the peak SAR values occurring in the head and hand for each of the configurations in Fig. 12 and 13. The two sets of data correspond to the configurations where the handset is rotated 60° and the handset is upright with respect to the head. As can be seen, each of the antenna structures results in very similar values of the peak SAR for both handset orientations with the exception of the side-mounted PIFA geometry. In this case, the rotated handset places the antenna nearly in contact with the ear tissue, resulting in a higher peak SAR as compared to the result for the upright handset. For all of the configurations, the peak

SAR in the head occurs either in the ear tissue or in the skin/fat layer in the antenna vicinity. Table 2 also provides the SAR averaged over the entire head, which as expected is considerably lower than the peak SAR levels. The numbers in Table 2 and plots in Figs. 12 and 13 show that while the spherical model is useful for obtaining a first-order approximation to the effect of the head, an inhomogeneous model is required to obtain the detailed field variations in and around the tissue. It is important that the numbers provided here are consistent with previously reported measured and computed results [7], [8], [18]. Comparisons such as these are very useful for determining the suitability of different radiators for personal communications applications.

E. Back-Mounted PIFA

In an effort to reduce the exposure to the head, a PIFA element is mounted on the back of the handset, as shown in Fig. 1(d). Since the conducting handset chassis lies between the operator and the antenna, it is expected that the peak SAR in the head will be reduced. The resulting field and SAR plots are provided in Figs. 12 and 13, respectively. These results, coupled with the numerical results in Table 2, serve to illustrate the substantial reduction in the peak SAR resulting from judicious placement of the element. Further refinement of similar antenna configurations may be possible to reduce the peak SAR and power absorption within the body.

The preceding results for radiation efficiency and peak SAR in the head have all been provided for a given separation between the head and the handset (b = 1.97 cm). However, it is interesting to examine the effect of this distance on these parameters. Fig. 14 presents the variation of the antenna efficiency and peak SAR (1 W delivered power) in the head versus the distance b for the monopole and the back-mounted PIFA configurations. As might be expected, the radiation efficiency increases with distance, while the peak SAR decreases in a nearly exponential fashion. It is noteworthy that as b increases, the back-mounted PIFA efficiency actually falls below that of the monopole. This phenomenon occurs because at these larger distances, the absorption in the hand becomes a more significant percentage of the total absorption. Since the power absorption in the hand induced by the PIFA is greater than that caused by the monopole (see Table 2), the PIFA efficiency is expected to fall below that of the monopole for large distances.

IV. Conclusion

This paper has presented some key issues relating to the electromagnetic interaction between handset-mounted antennas and a human operator's biological tissue. The study has used the FDTD simulation technique in conjunction with detailed models of the antennas, handset, head, and hand. The FDTD implementation summarized in the paper has been made adequately flexible to accommodate modeling of a wide variety of antenna and tissue topologies.

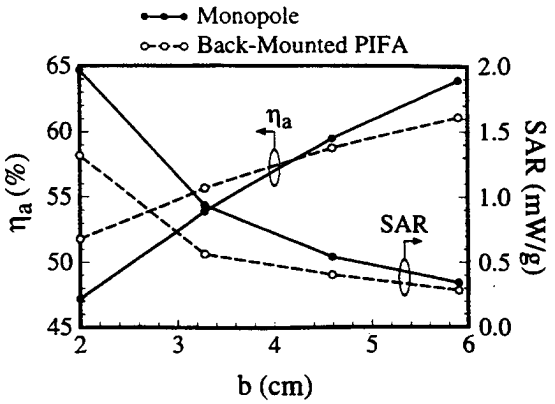


Fig. 14. Radiation efficiency η_a (%) and peak SAR in the head (P_{del} = 1 W) for the monopole and back-mounted PIFA versus the distance b between the head and handset (d = 7.21 cm).

The simulations presented for representative handset/tissue geometries have revealed that the tissue exercises a noticeable effect on the antenna input impedance, radiation patterns, and gain for both external and internal configurations. Most particularly, masking of low-profile antennas by the hand can result in serious detuning of the antenna impedance. In many cases, experimental results obtained at very low powers have been presented which verify the simulation accuracy and provide additional understanding into the antenna behavior. The numerical simulations also reveal that the SAR in the head has a peak between 2.0 and 3.8 mW/g and an average between 0.08 and 0.10 mW/g for 1 W of delivered power, and that the tissues absorb between 53 and 68% of the power delivered to the antenna for a head-handset separation of 2 cm. It has been shown that by placing an internal antenna (PIFA) on the back of the handset away from the user, the peak and average SAR in the head and the power absorption may be reduced to 0.9 mW/g, 0.06 mW/g, and 48%, respectively.

References

[1] J. J. C. Chang, R. A. Miska, and R. A. Shober, "Wireless systems and technologies: an overview," *AT&T Tech. J.*, vol 72, pp. 11–18, July/Aug. 1993.

[2] K. S. Yee, "Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media," *IEEE Trans. Ant. Propag.*, vol. AP-14, pp. 302–307, May 1966.

[3] A. Taflove and M. E. Brodwin, "Numerical solution of steady-state electromagnetic scattering problems using the time-dependent Maxwell's equations," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-23, pp. 623–630, Aug. 1975.

[4] R. Luebbers, L. Chen, T. Uno, and S. Adachi, "FDTD calculation of radiation patterns, impedance, and gain for a monopole antenna on a conducting box," *IEEE Trans. Ant. Propag.*, vol 40, pp. 1577–1583, Dec. 1992.

[5] M. A. Jensen and Y. Rahmat-Samii, "Performance analysis of antennas for hand-held transceivers using FDTD," *IEEE Trans. Ant. Propag.*, vol. 42, pp. 1106–1113, Aug. 1994.

[6] M. A. Jensen and Y. Rahmat-Samii, "FDTD analysis of PIFA diversity antennas on a hand-held transceiver unit," in *1994 IEEE AP-S Int. Symp. Dig.*, vol. 2, pp. 814–817, Ann Arbor MI, June/July 1993.

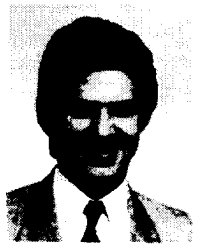
[7] J. Toftgård, S. N. Hornsleth, and J. B. Andersen, "Effects on portable antennas of the presence of a person," *IEEE Trans. Ant. Propag.*, vol. 41, pp. 739–746, June 1993.

[8] P. J. Dimbylow, "FDTD calculations of the SAR for a dipole closely coupled to the head at 900 MHz and 1.9 GHz," *Phys. Med. Biol.*, vol. 38, pp. 361–368, Feb. 1993.
[9] P. J. Dimbylow and O. P. Gandhi, "Finite-difference time-domain calculations of SAR in a realistic heterogeneous model of the head for plane-wave exposure from 600 MHz to 3 GHz," *Phys. Med. Biol.*, vol. 36, pp. 1075–1089, Aug. 1991.
[10] D. M. Sullivan, O. P. Gandhi, and A. Taflove, "Use of the finite-difference time-domain method for calculating *EM* absorption in man models," *IEEE Trans. Biomed. Eng.*, vol. 35, pp. 179–186, Mar. 1988.
[11] G. Mur, "Absorbing boundary conditions for the finite-difference approximation of the time-domain electromagnetic field equations," *IEEE Trans. Electromagnetic Compat.*, vol. EMC-23, pp. 377–382, Nov. 1981.
[12] A. C. Eycleshymer and D. M. Schoemaker, *A Cross-Section Anatomy*. New York: Appleton-Century, 1911.
[13] D. M. Sullivan, "A frequency-dependent FDTD method for biological applications," *IEEE Trans. Microwave Theory Tech.*, vol. 40, pp. 532–539, Mar. 1992.
[14] American National Standards Institute, IEEE C95.1–1991: *IEEE Standard for Safety Levels with Respect to Human Exposure to Radio Frequency Electromagnetic Fields, 3 kHz to 300 GHz*, 1991, Piscataway, NJ 08855–1331.
[15] M. Fischetti, "The cellular phone scare," *IEEE Spectrum*, pp. 43–47, Mar. 1993.
[16] R. J. Luebbers and K. Kunz, "FDTD modeling of thin impedance sheets," *IEEE Trans. Ant. Propag.*, vol. 40, pp. 349–351, Mar. 1992.
[17] T. Taga and K. Tsunekawa, "Performance analysis of a built-in planar inverted *F* antenna for 800 MHz band portable radio units," *IEEE J. Selected Areas Commun.*, vol. SAC-35, pp. 921–929, June 1987.
[18] H. R. Chuang, "Human operator coupling effects on radiation characteristics of a portable communication dipole antenna," *IEEE Trans. Ant. Propag.*, vol. 42, pp. 556–560, Apr. 1994.

Michael A. Jensen (Student Member, IEEE) received the B.S. (summa cum laude) and M.S. degrees in electrical engineering from Brigham Young University (BYU) in 1990 and 1991, respectively. He is currently pursuing the Ph.D. in electrical engineering at the University of California, Los Angeles.

From 1989 to 1991 he was a graduate Research Assistant in the Lasers and Optics Laboratory at BYU. Since 1991, he has been with the Antenna Laboratory at UCLA as a graduate student researcher. His main research interests include radiation and propagation for personal communications, numerical electromagnetics, optical fiber communication, and implementation of finite-difference schemes on massively parallel computer architectures.

Mr. Jensen is a member of Eta Kappa Nu and Tau Beta Pi. In 1990 he received a National Science Foundation Graduate Fellowship.



Yahya Rahmat-Samii (Fellow, IEEE) received the M.S. and Ph.D. degrees in electrical engineering from the University of Illinois at Champaign–Urbana.

He is a Professor of Electrical Engineering at the University of California, Los Angeles, and was a Guest Professor at the Technical University of Denmark (TUD) in 1986. He has also been a Senior Research Scientist at the NASA Jet Propulsion Laboratory/California Institute of Technology since 1978, where he has contributed to the advancement of antenna technology for space programs. He has made pioneering contributions to the developments of near-field plane-polar and bipolar antenna measurements, microwave holographic diagnostics, mobile satellite communication antennas, reflector surface compensation, multireflector antenna diffraction analysis and synthesis, scattering and radiation from complex objects, RCS computations, singularity in dyadic Green's function, high power microwave (HPM) antennas, EMP and aperture penetration, the spectral theory of diffraction (STD), and GTD. He has authored or coauthored over 300 technical journal articles and conference papers and has written chapters in 13 books.

Dr. Rahmat-Samii became a Fellow of IAE in 1986 and was the 1984 recipient of the Henry Booker Award of URSI. He was appointed an IEEE Antennas and Propagation Society Distinguished Lecturer and presented lectures internationally. He was an elected IEEE AP-S AdCom member for the second term and has been an Associate Editor of the *IEEE Transactions on Antennas and Propagation* and the society's *Magazine*. He is currently the elected President of IEEE AP-S. He was the Chairman of the IEEE Antennas and Propagation Society of Los Angeles during 1987–1989, which received the Antennas and Propagation Best Chapter Award from the AP Society in 1989. He is one of the three international Editors of the IEE book series on electromagnetics and antennas. He is also one of the Editors of the *Journal of Electromagnetic Waves and Applications*. He is a Director of AMTA (Antenna Measurements Technique Association) and the Electromagnetics Society. He is listed in *Who's Who in America*, *Who's Who in Frontiers of Science and Technology*, and *Who's Who in Engineering*. He has received numerous NASA Certificate of Recognitions and recently earned the JPL Team NASA's Distinguished Group Achievement Award. In 1992, he was the recipient of the Best Application Paper Award (Wheeler Award) for a paper published in the *Transactions* in 1991. He is a member of Commissions A, B, and J of USNC/URSI, Sigma Xi, Eta Kappa Nu, and the Electromagnetics Academy.

Channel Coding Strategies for Cellular Radio

G.J. Pottie and A.R. Calderbank

IEEE Transactions on Vehicular Technology
vol. 44, no. 4, pp. 763-770
April 1995.

Channel Coding Strategies for Cellular Radio

Gregory J. Pottie, *Member, IEEE,* and A. Robert Calderbank, *Member, IEEE*

***Abstract*—To improve re-use of time/frequency slots in a cellular radio system, it is desirable for the average interference levels seen by all users to be made approximately equal. We provide constructions based on orthogonal latin squares that guarantee different sets of users will interfere in successive slots. When signal to interference ratios are independent from successive symbols, channel codes can provide a large diversity gain which is far in excess of the gain against additive noise. Consequently, coding with interleaving fits together very naturally with this construction. We illustrate how to achieve large performance improvement using convolutional codes with low decoding delay in a slow hopped system.**

I. Introduction

In MOBILE cellular radio, the dominant impairments are multipath fading and interference from other mobiles. In a conventional time division multiple access system (TDMA), mobiles are assigned slots which they keep from frame to frame, with the interfering mobiles assigned slots in the same way. As interference levels vary widely between slots due to such factors as shadowing and geographic proximity, some mobiles suffer from persistently poor signal to interference ratios (which we shall refer to as C/I). Systems are generally designed for 90 and 99% worst case conditions. Therefore, the result of this uneven interference distribution is overly conservative restrictions on frequency re-use between cells, and thus reduced capacity.

If instead slot assignments were arranged such that different interferers were encountered in successive frames or slots and repetition or channel coding used to provide averaging, then the worst case error statistics would improve. In [1] random hopping over eight frequencies was used to average over different interferers and channel conditions. As a result, all frequencies could be re-used between cells. In [2], a combination of coding, explicit frequency re-use restrictions, and slow frequency hopping was employed. Both yielded capacities similar to that of the direct sequence spread spectrum code division multiple access (CDMA) system analyzed in [3]. Here we provide a specific construction of the hopping patterns which when combined with convolutional coding leads to improved performance, without restrictions on frequency re-use between cells. The construction essentially randomizes the signal to interference ratios encountered in successive hops. Channel codes in combination with interleaving may then be used to achieve a large diversity gain in addition to the coding gain against additive white Gaussian noise. Thus, hopping for interference diversity and channel coding are a natural match. Performance for our slow-hopped example is similar to that obtained in [4], but with lesser complexity. In that system, long rate-1/2 convolutional codes and hopping patterns with low cross-correlation were employed. In Section II we describe our model of the propagation environment and discuss the need for randomization of interferers in TDMA or hybrid frequency hopped TDMA. In Section III we describe our construction, which is based on orthogonal latin squares. We discuss the basic coding tradeoffs for hopped systems using latin squares or similar techniques in Section IV, and illustrate them with two computer simulated examples. We demonstrate significant gain with low delay (20–40 ms) and modest complexity using convolutional codes. In Section V we present our conclusions.

Manuscript received January 15, 1993; revised April 3, 1994. This work supported in part by AT&T Bell Laboratories and ARPA contract DAAB07-93-C-C501.

G. J. Pottie is with the Electrical Engineering Department, University of California, Los Angeles, CA 90024 USA.

A. R. Calderbank is with the Mathematical Sciences Research Center, AT&T Bell Laboratories, Murray Hill, NJ 07974 USA.

IEEE Log Number 9411304.

II. The Interference Environment

We follow the simplified propagation model presented in [3], and focus on the reverse or up-link of mobile to base station. This choice is arbitrary for synchronous systems, but assists in comparison of results in the literature. Mobiles are assumed to be uniformly distributed on an infinite plane, with base stations arranged in a uniform hexagonal pattern. Power drops with the fourth power of distance. In addition, there is independent log-normal shadowing for each track. The received power level for a mobile m at distance d from a base station is then

$$P = \frac{10^{0.1\xi} P_m}{d^4}$$

where P_m is the mobile's transmit power, and ξ is a Gaussian random variable with zero mean and standard deviation $\sigma = 8$. The mobiles are assigned to the base station with the minimum propagation loss.

One practical but suboptimal power control strategy is to attempt to make the power level received at the base station approximately the same for all mobiles in the cell. The specific level depends on the margin required against additive white Gaussian noise, but we shall assume the power level to be unity, since all powers and interference levels will scale. Let cell 0 be our reference cell, i.e., the set of mobiles communicating with base station 0. Then the interference power encountered at the base station of cell 0 from a mobile belonging to cell i is

$$I = \left(\frac{r_1}{r_0}\right)^4 10^{0.1(\xi_0 - \xi_i)}$$

where

r_i distance of mobile to base station i.
r_0 distance of mobile to base station 0,
ξ_0, ξ_i shadowing random variables for tracks to base stations 0 and i, respectively.

Shadowing can be assumed to be constant over the short run. With this propagation model and channel assignment strategy, most of the interference encountered in cell 0 is generated in the two surrounding rings of cells.

Shadowing significantly reduces the geographic isolation of cells. Computations performed over the first two complete rings with only fourth power propagation loss and with power controlled for equal received levels at the home base stations indicate that a regular frequency re-use pattern of 1/3 improves C/I by close to 13 dB compared to use of all frequencies in all cells. However, when shadowing is included in the model the improvement shrinks to 7 dB. This compares to a gain of 4.8 dB which might be obtained just by reducing the density of users in all cells by this same fraction. This may be intuitively explained as follows. With shadowing, the hexagonal cell geometry breaks down and some mobiles may wander far outside these nominal boundaries. In particular, mobiles belonging to the second ring of cells may wander much closer to the base station for cell 0. At the same time, the interference power due to such mobiles is upper bounded by 1, since otherwise they would be handed off to cell 0. With only attenuation with distance, most interference power originated in the region of the first ring of cells near the boundary with cell 0, with interference power near 1. Shadowing merely smears out this region. Thus, the probability of large interference power increases significantly for mobiles assigned to the second ring of cells, while not increasing at all for the first ring of cells. The end result is that it does not matter much for the average C/I whether a reduction in mobile density is achieved through explicit frequency re-use patterns or through limits on numbers within each cell, with full frequency re-use. The advantage of the latter approach is that the limits can easily be made adaptive, and no frequency use planning is required.

For a system with frequency re-use of 1, by symmetry, whether one assumes there to be three perfect 120 sectors or ordinary cells with 1/3 the density of mobiles the interference levels will be the same. Consequently, all our computations will be performed assuming no sectorization; the only normalization required to directly compare our results to a sectorized system is to multiply the number of mobiles by 3.

Our simulations assume non-standard narrowband (<20 kHz) systems. Consequently, a reasonable worst-case model for the short term channel conditions is Rayleigh fading, i.e., the signal is multiplied by a complex Gaussian, with variance in each dimension equal to half the computed long-term power. We suppose that transmission is divided into time or frequency slots. Time slots may hop in position within the frame or in frequency from frame to frame. Usually we assume that fading is independent from slot to slot for any given user, which may be achieved by some combination of delay and frequency agility. We further assume that the slots are short, so that the signal level is constant over the slot (including Rayleigh fading).

In such a system, if mobiles in nearby cells share the same pattern of frequency hops then diversity is obtained against multipath fading in both the desired signal and interference. However, because of the long-term propagation conditions outlined above, the average C/I will vary widely between different mobiles, resulting in poor 90 and 99% worst case statistics. If coding across the slots were employed, and different sets of interferers were encountered in successive slots, then a channel code would provide diversity protection against these long-term interference levels, improving the error performance. In addition, there would be less variation in the error performance of the users. In the next section, we outline how to achieve maximum interferer diversity for a given frame length, while in Section IV we quantify the performance advantage obtained with such schemes.

III. Latin Squares

A latin square of order n is an $n \times n$ matrix with entries from a set R of n distinct elements, say $R = \{0, 1, \cdots, n-1\}$ such that each row and column contains every element of R exactly once. The entries in R might represent different users in the same cell, the rows representing different frequency slots, and the columns different time slots. Consider for example the matrix labelled {1} that appears in Fig. 1: in successive time slots 0, 1, 2, 3, 4 user 1 is assigned frequency slots 1, 0, 4, 3, 2, respectively. Two latin squares are said to be orthogonal if the n^2 ordered pairs (i, j), where i and j are the entries from the same position in the respective squares, exhaust the n^2 possibilities, i.e., every ordered pair occurs exactly once. Orthogonality of latin squares corresponds to there being exactly one time/frequency collision for every pair of users in different cells. For example, user 3 in matrix {2} is assigned frequencies 4, 1, 3, 0, 2 in successive slots, which overlap user 1 of {1} only at time 4.

If n is prime, then there is a simple construction for a family of n-1 mutually orthogonal latin squares. For $a = 1, \cdots, n-1$ we define an $n \times n$ matrix $\{a\}$ by setting

$$\{a\}_{ij} \equiv ai + j (\text{mod } n).$$

It is easy to verify that each matrix is a latin square: for fixed i, the numbers $ai + j, j = 0, \cdots, n-1$ are distinct, and for fixed j, the numbers $ai + j, i = 0, \cdots, n-1$ are distinct, in both cases assuming modulo arithmetic. It remains to verify that if $a \neq b$, then the matrices $\{a\}$ and $\{b\}$ are orthogonal. This amounts to showing that for all $c, d = 0, 1, \cdots, n-1$, there is a unique solution (i, j) to the linear system

$$c \equiv ai + j$$
$$d \equiv bi + j.$$

This follows from the fact that the determinant of this linear system is $a - b$, which is non-zero (mod n).

If n is prime, then the residues modulo n form the finite field of order n. The above construction extends to finite fields of prime power order in the natural way: now R is the set of field elements, and for every non-zero field element a, we

$$\{1\} = \begin{bmatrix} 0 & 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 & 0 \\ 2 & 3 & 4 & 0 & 1 \\ 3 & 4 & 0 & 1 & 2 \\ 4 & 0 & 1 & 2 & 3 \end{bmatrix} \quad \begin{matrix} \text{rows} \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{matrix}$$

$$\{2\} = \begin{bmatrix} 0 & 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 0 & 1 \\ 4 & 0 & 1 & 2 & 3 \\ 1 & 2 & 3 & 4 & 0 \\ 3 & 4 & 0 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 & 0 \\ 2 & 3 & 4 & 0 & 1 \\ 3 & 4 & 0 & 1 & 2 \\ 4 & 0 & 1 & 2 & 3 \end{bmatrix}$$ ← permutation P_2

$$\{3\} = \begin{bmatrix} 0 & 1 & 2 & 3 & 4 \\ 3 & 4 & 0 & 1 & 2 \\ 1 & 2 & 3 & 4 & 0 \\ 4 & 0 & 1 & 2 & 3 \\ 2 & 3 & 4 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 & 0 \\ 2 & 3 & 4 & 0 & 1 \\ 3 & 4 & 0 & 1 & 2 \\ 4 & 0 & 1 & 2 & 3 \end{bmatrix}$$ ← permutation P_3

$$\{4\} = \begin{bmatrix} 0 & 1 & 2 & 3 & 4 \\ 4 & 0 & 1 & 2 & 3 \\ 3 & 4 & 0 & 1 & 2 \\ 2 & 3 & 4 & 0 & 1 \\ 1 & 2 & 3 & 4 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 & 0 \\ 2 & 3 & 4 & 0 & 1 \\ 3 & 4 & 0 & 1 & 2 \\ 4 & 0 & 1 & 2 & 3 \end{bmatrix}$$ ← permutation P_4

Fig. 1. Four mutually orthogonal latin squares of order 5.

define a matrix $\{a\}$ with the ijth entry $\{a\}_{ij} = ai + j$, where addition and multiplication take place in the field.

There is a rich combinatorics literature that relates mutually orthogonal latin squares, projective planes, and planar ternary rings. We refer the reader to Hughes and Piper [5] for more information.

A. Examples

For $n = 5$, the construction described above produces 4 mutually orthogonal latin squares, as listed in Fig. 1. Notice that the matrices $\{2\}, \{3\}, \{4\}$ can all be obtained from $\{1\}$ by permuting rows. These row permutations are all powers of P^2: $P_4 = P_2^2$, and $P_3 = P_2^3$ which follows from $4 \equiv 2^2 \pmod 5$ and $3 \equiv 2^3 \pmod 5$.

For $n = 8$, the construction makes use of the finite field F_8 which results from appending a root of the primitive polynomial $p(x) = x^3 + x + 1$ to $\{0, 1\}$. The elements of the field, and the resulting matrices and listed in Fig. 2. Again note that every matrix is obtained from $\{1\}$ by permuting rows using some power of the permutation P_2. This will be true in general because every field has a primitive element. In this example, 2 represents x, and the powers of x exhaust the non-zero field elements.

B. Implementation

The above construction may be used in a synchronous cellular radio system. That is, the timing of the slot boundaries must be aligned across cells, with guard bands potentially required for large cells to account for propagation delays from nearby cells. Each cell is assigned a matrix $\{a\}$ in the set of $n-1$ mutually orthogonal latin squares. Then up to n users can be accommodated in any cell, each user being assigned a field element. For example, suppose $n = 8$, $a = 5$, and a user is assigned the element 7. As described previously, we interpret

$$\{1\} = \left[\begin{array}{cccc|cccc} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 1 & 0 & 4 & 7 & 2 & 6 & 5 & 3 \\ 2 & 4 & 0 & 5 & 1 & 3 & 7 & 6 \\ 3 & 7 & 5 & 0 & 6 & 2 & 4 & 1 \\ \hline 4 & 2 & 1 & 6 & 0 & 7 & 3 & 5 \\ 5 & 6 & 3 & 2 & 7 & 0 & 1 & 4 \\ 6 & 5 & 7 & 4 & 3 & 1 & 0 & 2 \\ 7 & 3 & 6 & 1 & 5 & 4 & 2 & 0 \end{array}\right] \quad \begin{matrix} \text{rows} \\ 0=0 \\ 1=1 \\ 2=x \\ 3=x^2 \\ 4=x+1 \\ 5=x^2+x \\ 6=x^2+x+1 \\ 7=x^2+1 \end{matrix}$$

$$\{2\} = \left[\begin{array}{cccc|cccc} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 2 & 4 & 0 & 5 & 1 & 3 & 7 & 6 \\ 3 & 7 & 5 & 0 & 6 & 2 & 4 & 1 \\ 4 & 2 & 1 & 6 & 0 & 7 & 3 & 5 \\ \hline 5 & 6 & 3 & 2 & 7 & 0 & 1 & 4 \\ 6 & 5 & 7 & 4 & 3 & 1 & 0 & 2 \\ 7 & 3 & 6 & 1 & 5 & 4 & 2 & 0 \\ 1 & 0 & 4 & 7 & 2 & 6 & 5 & 3 \end{array}\right] = \left[\begin{array}{cccc|cccc} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 1 & 0 & 4 & 7 & 2 & 6 & 5 & 3 \\ 2 & 4 & 0 & 5 & 1 & 3 & 7 & 6 \\ 3 & 7 & 5 & 0 & 6 & 2 & 4 & 1 \\ \hline 4 & 2 & 1 & 6 & 0 & 7 & 3 & 5 \\ 5 & 6 & 3 & 2 & 7 & 0 & 1 & 4 \\ 6 & 5 & 7 & 4 & 3 & 1 & 0 & 2 \\ 7 & 3 & 6 & 1 & 5 & 4 & 2 & 0 \end{array}\right] \quad P_2$$

$$\{3\} = \left[\begin{array}{cccc|cccc} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 3 & 7 & 5 & 0 & 6 & 2 & 4 & 1 \\ 4 & 2 & 1 & 6 & 0 & 7 & 3 & 5 \\ 5 & 6 & 3 & 2 & 7 & 0 & 1 & 4 \\ \hline 6 & 5 & 7 & 4 & 3 & 1 & 0 & 2 \\ 7 & 3 & 6 & 1 & 5 & 4 & 2 & 0 \\ 1 & 0 & 4 & 7 & 2 & 6 & 5 & 3 \\ 2 & 4 & 0 & 5 & 1 & 3 & 7 & 6 \end{array}\right] = \{1\}P_2^2 \qquad \{4\} = \left[\begin{array}{cccc|cccc} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 4 & 2 & 1 & 6 & 0 & 7 & 3 & 5 \\ 5 & 6 & 3 & 2 & 7 & 0 & 1 & 4 \\ 6 & 5 & 7 & 4 & 3 & 1 & 0 & 2 \\ \hline 7 & 3 & 6 & 1 & 5 & 4 & 2 & 0 \\ 1 & 0 & 4 & 7 & 2 & 6 & 5 & 3 \\ 2 & 4 & 0 & 5 & 1 & 3 & 7 & 6 \\ 3 & 7 & 5 & 0 & 6 & 2 & 4 & 1 \end{array}\right] = \{1\}P_2^3$$

$$\{5\} = \left[\begin{array}{cccc|cccc} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 5 & 6 & 3 & 2 & 7 & 0 & 1 & 4 \\ 6 & 5 & 7 & 4 & 3 & 1 & 0 & 2 \\ 7 & 3 & 6 & 1 & 5 & 4 & 2 & 0 \\ \hline 1 & 0 & 4 & 7 & 2 & 6 & 5 & 3 \\ 2 & 4 & 0 & 5 & 1 & 3 & 7 & 6 \\ 3 & 7 & 5 & 0 & 6 & 2 & 4 & 1 \\ 4 & 2 & 1 & 6 & 0 & 7 & 3 & 5 \end{array}\right] = \{1\}P_2^4 \qquad \{6\} = \left[\begin{array}{cccc|cccc} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 6 & 5 & 7 & 4 & 3 & 1 & 0 & 2 \\ 7 & 3 & 6 & 1 & 5 & 4 & 2 & 0 \\ 1 & 0 & 4 & 7 & 2 & 6 & 5 & 3 \\ \hline 2 & 4 & 0 & 5 & 1 & 3 & 7 & 6 \\ 3 & 7 & 5 & 0 & 6 & 2 & 4 & 1 \\ 4 & 2 & 1 & 6 & 0 & 7 & 3 & 5 \\ 5 & 6 & 3 & 2 & 7 & 0 & 1 & 4 \end{array}\right] = \{1\}P_2^5$$

$$\{7\} = \left[\begin{array}{cccc|cccc} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 7 & 3 & 6 & 1 & 5 & 4 & 2 & 0 \\ 1 & 0 & 4 & 7 & 2 & 6 & 5 & 3 \\ 2 & 4 & 0 & 5 & 1 & 3 & 7 & 6 \\ \hline 3 & 7 & 5 & 0 & 6 & 2 & 4 & 1 \\ 4 & 2 & 1 & 6 & 0 & 7 & 3 & 5 \\ 5 & 6 & 3 & 2 & 7 & 0 & 1 & 4 \\ 6 & 5 & 7 & 4 & 3 & 1 & 0 & 2 \end{array}\right] = \{1\}P_2^7$$

Fig. 2. Seven orthogonal latin squares of order 8.

the rows as frequency slots and the columns as time slots for a frequency hopped system. Then at time 0, frequency 3 is transmitted, at time 1 frequency 6 is transmitted, etc. Except for other cells which also are assigned matrix $\{5\}$ there will be at most one collision with any particular user at any given time/frequency pair over the 8 slot duration of the frame. If a regular re-use pattern of 7 cells is used for the matrices, then the signals from users sharing the same hopping patterns will be heavily attenuated. In this way, each user suffers collisions from an ensemble of users occupying the nearby cells, rather than with one dominant interferer. Obviously, with a larger value of n, the re-use distance increases and the randomization of interference hits is more complete. Also, the rows might be reinterpreted as time slots in a frame, and the columns as frame positions within a superframe for a pure TDMA system.

The construction allows each mobile to quickly generate its own time/frequency hopping pattern given minimal information transfer from the base station. The only information required, given that n and the primitive polynomial $p(x)$ are known, is the matrix number, a, and the field element, e. Then if i is the row number, the column corresponding to i is just $j \equiv e + ai \pmod{p(x)}$. Hence, the overhead involved in communicating the hopping pattern is negligible. Note that we cannot get every possible value of n using this construction. However, primes and prime powers are dense for moderate n, affording a wide range of choices.

A very similar construction has recently been proposed by Wyner [6] using a combinatorial configuration constructed by Roche [7] for use in wideband multi-tone systems, to assure that users in nearby cells interfere in at most one tone.

IV. Coding Strategies

There are several ways in which orthogonal latin squares may be used in a wireless system to improve capacity. The straightforward approach is to use repetition coding with some combining strategy such as maximal ratio combining or slot selection. Then every user will experience the average interference level, or close to it, which represents a large improvement over the worst case conditions. However, it is very wasteful of bandwidth. A more fruitful approach is to use channel coding in combination with interleaving to spread information over multiple slots. Channel codes provide gain in this situation in two ways. First, they provide the usual coding gain, i.e., for a particular information rate, a probability of error specification can be met with a lower signal to interference ratio. More significantly, provided the signal to interference ratios for successive symbols encountered by the decoder are uncorrelated, the diversity order of the system is increased, changing the shape of the error curve with respect to average signal to interference ratio. The diversity order obtained with coding is directly related to the minimum distance of the code, and the manner in which it is decoded. With the proper choice of decoding metric, the channel coder effectively performs diversity combining to find the most likely transmitted sequence. Thus, the latin squares technique forms a natural match to coding and interleaving, with the combination yielding both coding and diversity gain.

The use of channel coding with the latin squares technique results in an interference-limited system. That is, C/I is determined by the fraction of slots occupied, rather than by explicit spatial frequency re-use restrictions. In the following, we outline the consequences of this for channel code selection, and then discuss some of the trade-offs for channel code design, code rate, slot duration, and metric choice for a low-delay system. These choices result in significant performance differences. Among the many possible channel coding schemes that meet the low-delay restriction, we confine discussion to those that make use of convolutional encoders. The section is concluded with two illustrative examples.

A. Trellis Coding Alternatives

In conventional convolutional coding, every bit is encoded and extra symbols are transmitted to accommodate the redundancy introduced by the encoder. If the same energy is allocated to each information bit transmitted, the interference power is not changed in the system. Nevertheless, the usual bandwidth expansion penalty is paid against AWGN, and there is a price in signal to interference ratio, as follows. Consider a rate-1/2 code. An uncoded system with the same throughput, bandwidth, and energy would use half the slots with twice the power, and produce the same interference. Thus, the redundancy introduced in coding degrades the signal to interference ratio by 3 dB as for additive white noise. However, because there is coding gain, the coded system can actually operate at a much lower power level, or alternatively tolerate an increased level of interference, i.e., allow more users.

In trellis coded modulation the redundancy of the encoder is accommodated in expansion of the signal constellation. For example, rather than sending 2 bits per symbols using uncoded 4-PSK, we might use a rate-2/3 code and 8-PSK. Clearly both bandwidth and interference are unaffected by such schemes. However, the distance between signal points in the signal constellation shrinks, resulting in a power penalty against both noise and interference. This penalty is much more severe than that obtained with ordinary convolutional coding, and should be paid only if bandwidth expansion is not possible, as for example in TDMA systems. Define the efficiency η as the time-bandwidth product occupied by active mobiles divided by the time-bandwidth product of the system. For the propagation environment outlined in Section II, a system with $\eta = 1$ would have a signal to interference ratio of 3 dB. Clearly the number of users must be drastically reduced to permit an uncoded system to operate. On the other hand, coded systems can operate at very low signal to noise ratios. Rate $1/n$ codes can successfully operate at lower SNR with the same number of states as n gets larger, although the coding gain against AWGN (which includes the rate-reduction penalty) is to first order independent of n (see [8]). For the same energy per transmitted bit, use of a rate $1/n$ code improves the signal to interference ratio per bit by the factor of n, the same effect as reducing the occupancy in an uncoded system by the factor n. However, channel codes provide gain, and so presuming that there remain open slots, the number of users can actually be increased relative to the uncoded system. This tradeoff is different from bandwidth-limited systems, where use of convolutional coding implies a reduction in the number of users, if the frequency re-use is not changed. Thus, convolutional codes provide a more favorable gain/complexity tradeoff in this application than TCM, and increase the number of users permitted to be simultaneously active.

It remains for us to address the question of how to use convolutional codes to obtain coding gain and diversity protection with moderate decoding delay and complexity. To make matters more concrete, we assume that a slow-hopped system employing differential quadrature phase shift keying (DQPSK) is used. The intended application is compressed speech at 8 kb/s. An uncoded system might be operated at 4 kbaud, occupying a bandwidth of 5 kHz. The symbol duration is 250 μs, which is very large compared to either the delay spread or group delay variation of the typical mobile channel. A convolutional code will be used, with interleaving across slots. In 20 ms, 80 symbols are transmitted. To achieve reasonable frequency diversity in 20 ms and keep the decoding and interleaving delay reasonable, we must hop every few symbols.

For coding to realize the maximum interference and frequency diversity, the signal to interference ratio for successive symbols should be uncorrelated. This implies that coding must be interleaved across slots. For a delay constraint of 20 ms and slot duration of m ms, the truncation depth of the Viterbi decoder can be no more than 20/m. Indeed, it can be less if each code branch corresponds to several channel symbols. For example suppose we use a rate-1/2 code for 8 kb/s speech,

and modulate using DQPSK at 8 kbaud. If slots consist of 8 symbols, then each slot lasts 1 ms and a truncation depth of 20 is permitted. The full minimum distance of the 32 state code is obtained with a truncation depth of 19 [8], while codes with larger minimum distance will not provide significantly better performance unless the truncation depth is increased. Thus, complexity is limited by delay unless shorter slots are used. An alternative to the rate-1/2 code is the 32 state rate-1/4 code with minimum distance 18. Keeping 8-symbol slots but now signalling at 16 kbaud, there are a total of 320 symbols in 20 ms, or 40 slots. Each code branch corresponds to symbols in two successive slots. Therefore by decreasing the rate of the code, diversity can be increased. There is little cost in complexity since there are still the same number of states and branches entering each state, and branches are processed at the same speed. Moreover, the truncation depth in branches is a less severe function of minimum distance for rate-1/4 codes [9], so that roughly the same depth versus number of states tradeoff is obtained. On the other hand, the branch computations are more complicated since demodulation must be accomplished at twice the speed, and there is less flexibility in increasing the number of mobiles when the interference from nearby cells is lighter. Rate-1/2 and rate-1/4 codes will be compared in the simulated examples.

B. Decoding Metrics

Any diversity combining device benefits from channel state information, i.e., knowledge of the signal and interference levels. Accurate information is quite difficult to obtain in mobile radio. If short slots are used, the interference estimate will be poor; if longer slots are used, the signal and interference statistics change over the slot duration. We now consider how to estimate the channel state information, and how the quality of the estimates influence the choice of decoder metric. We start with neither a phase reference nor knowledge of the signal power. Signal power may be estimated as the average power over the slot. The mean squared error may be obtained by computing the minimum distance between any pair of received points and the allowable phase transitions in the signal set. Let the set of received channel symbols over the slot be $\{r_i, i = 1, 2, \cdots, N\}$. Then the estimate of the signal power is

$$\hat{S} = \frac{1}{N} \sum_{i=1}^{N} |r_i|^2.$$

Let the set of transmitted signals be $\{x_k, k = 1, \cdots 4\}$ corresponding to four phases on the unit circle, with x_1 corresponding to a phase of 0. There are four allowable phase transitions in DQPSK between times i and $i+1$. The estimate of the squared error assuming phase transition k is given by

$$e_{ik} = \min_{\theta} \{(r_{ni} - x_1 e^{j\theta})^2 + (r_{ni+1} - x_k e^{j\theta})^2\}$$

where r_{ni} is the power-normalized received signal and the minimization is performed over all possible starting phases θ. The squared error estimate e_i for the phase transition between times i and $i+1$ is then found by minimizing over the x_k, so that the estimated mean squared error is given by

$$\hat{E} = \frac{1}{N-1} \sum_{i=1}^{N-1} e_i.$$

The estimate of C/I is then $\hat{S}/\hat{E}$. For example, consider a phase transition of 90 degrees. The reference points lie on the unit circle with a phase difference of 90 degrees, but with unknown starting angle. If this were the true transition, then the most likely starting angle would be the one that minimizes the squared distance between the (ordered) received sequences and the reference points. The distances can be tabulated for any combination of quantized and power-normalized received signal points for each allowable transition. The decision is for the transition with the least distance, and the estimate of the squared error is a value from the table. This procedure overestimates C/I because if errors actually occur we record the distance to the erroneous rather than the true symbol, and also because the total power includes the interference.

The C/I determined for the slot may be used to weight decisions. For example, it could be used to select a receive antenna or to weight the values for maximal ratio combining. It may also be used with a "soft" decoder metric. The squared distances for each allowable transition can be divided by $\hat{E}$ for the slot to form the branch metrics for a rate-1/2 code, or part of the branch metric for lower rates. This is not the optimal metric even if the C/I estimate is good; however the maximum likelihood metric requires a set of look-up tables to cover the range of C/I, which is very broad. This metric has the virtue of relative simplicity of implementation and includes both distance and signal to interference ratio information. It is in the spirit of a metric proposed by Ketchum and Wallace [10] for use with Reed-Solomon codes and QPSK, which while also not maximum likelihood, performed quite well in simulations.

We have conducted computer simulations of the 32-state rate-1/2 and rate-1/4 codes with this soft metric, and hard decision decoding with a slot erasure threshold (based on the estimate of C/I). We have considered two cases: 1) a benevolent genie supplies perfect C/I estimates and 2) C/I is estimated as above with a slot duration of 17 symbols. Representative results are presented in Table I. The simulations assume separate decoders, interleaved across slots. There are twenty slots per frame for rate-1/2 codes, and forty per frame for rate-1/4 codes. Here the interference hopping patterns were randomly generated with the provision that they hit users in cell 0 in only one slot per frame, roughly equivalent to the effect of the constructions of Section III. We see that the channel state information is extremely valuable, particularly in attempting to achieve low error rates. With perfect channel state information, soft decisions outperform hard decisions in all cases. However, the estimates obtained in practice are so poor that hard decision decoding with a 4 dB erasure threshold actually performs better for lower error rates for the rate-1/4 code.

Because the gap in performance between known and estimated C/I is so large we have investigated the possibility of estimating C/I from a known training sequence for each slot. Accounting for the overhead involved, the results are worse than for estimating it from the data. It would appear

TABLE I
EFFECT OF METRIC AND CHANNEL STATE INFORMATION ON ERROR PERFORMANCE

Code Rate	η	C/I (dB)	Probability of Error hard decoding, 4 dB C/I erasure threshold		Soft decoding	
			Perfect C/I knowledge	Estimated C/I	Perfect C/I knowledge	Estimated C/I
1/2	0.5	5.5	0.083	0.15	0.014	0.061
	0.25	8.3	0.014	0.029	2.1×10^{-4}	6.7×10^{-3}
	0.125	11.8	5.9×10^{-4}	3.5×10^{-3}	—	2.0×10^{-3}
1/4	1	2.3	0.010	0.14	1.9×10^{-3}	0.034
	.5	5.4	2.4×10^{-3}	5.4×10^{-3}	—	1.9×10^{-3}
	.25	8.4	4.3×10^{-5}	7.7×10^{-5}	—	8.9×10^{-4}

that a very large number of symbols are required to get a useful estimate of the interference power. Consequently, a hard decision metric is a reasonable choice, given the considerable reduction in decoder complexity. However, the relative merits of hard and soft decisions depend on the code, the antenna diversity scheme, and the rate of change of the channel.

C. Example 1

In this example, we consider DQPSK modulation. Each user gets 40 slots per frame, each time/frequency slot consisting of one reference symbol and 16 information-bearing symbols. The slots hop in frequency throughout the frame, and we assume that fading is independent in successive slots. Coding is across slots, accomplished using a rate-1/4 code, which selects two symbols in successive slots for each branch. Preceding the decoder there may be two-branch antenna selection diversity. It was found that performance was best using the criterion of the largest C/I estimate, rather than largest power estimate or combining using either of these criteria. Simulations were conducted over multiple user sets, for a variety of efficiencies. The figure of merit is the fraction of mobiles that experience a bit error rate in excess of 10^{-3} for a given use of time and bandwidth.

In Fig. 3 we plot the results for the assumption that antenna selection diversity is employed, and with users encountering the same interferers on each hop, labelled as curve 1. We plot call blocking probability versus number of 8 kb/s users per 1.25 MHz bandwidth, to assist in comparing to the CDMA results of [3]. We have removed the gain due to voice activity factor from the latter. All results assume equal cell loading. Note that 1.25 MHz is not wide enough for the assumption of independence of fading levels to be valid. For example, the CDMA calculations assume only diversity of order 2 for this bandwidth. Our system would have to hop over a wider bandwidth for the independence assumption to be valid, and obviously a direct sequence spread spectrum system could also make use of a wider bandwidth to improve diversity. It is clear in any case that performance of the slow-hopped system is inferior at the design target of 1% blocking probability. Thus, even though hopping with coding has produced frequency diversity, the large variations in average interference levels among users restricts the capacity.

However, the picture changes dramatically when a method such as latin squares is employed. In Fig. 3 we also plot results for single and dual antenna branch selection, in curves 2 and 3, respectively. Large gains are evident. Even without antenna

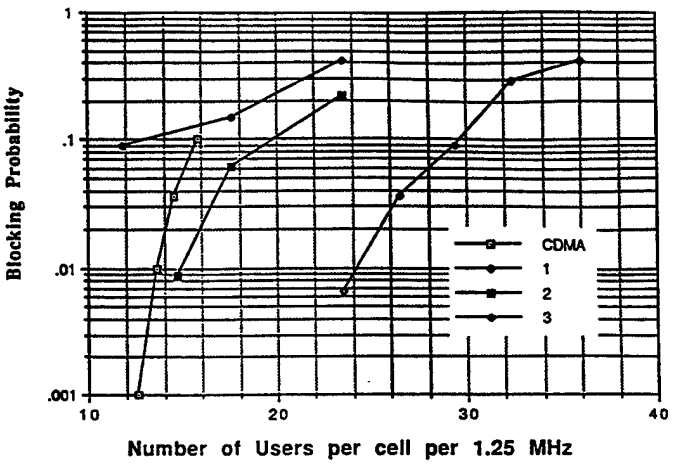


Fig. 3. Comparison of capacities.

diversity, the system gives roughly the same performance as DS-CDMA at the target of 1% blocking probability. The system with antenna diversity achieves nearly double the capacity. While it is true that this improvement is achieved assuming a wider hopping bandwidth, in addition our scheme has lower decoding delay and complexity.

Decoding delay in this example is roughly 40 ms. It can be reduced at some cost in efficiency by decreasing the number of symbols per slot to 8. The computer simulation results indicate roughly a 20% reduction in capacity for 1% blocking probability.

Also of interest is the increase in the number of users when the surrounding cells are more lightly loaded. In CDMA, the main limiting factor is the interference within the cell, which will cause the error rate to rise with the number of users. However, for a synchronous hopped system what limits the number of users in any given cell is the interference they cause in other cells. Consequently, the decision to increase the number of users must come from the network rather than through internal quality of service measurements. For example, if there are no calls in all the surrounding rings of cells, capacity increases to the limit of $\eta = 1$. However, even if the immediate neighbors have reduced loading, the allowable capacity increase depends on the loading of their neighbors, and the relative power we may assign to calls in the various cells. Thus, while this system shares with DS-CDMA the characteristic of "soft" capacity, optimum management of this capacity will be more intricate.

D. Example 2

In example 1, it was assumed that Rayleigh fading was independent in successive frequency hops. This presumes that hopping takes place over a very wide bandwidth. We now consider what happens when fading is static and wideband (relative to the hopping bandwidth). This might be the situation for example in a relatively narrowband microcellular system, with communication by pedestrians. Fading would in this case be very slow compared to the signaling interval, affording the receiver time to learn both the signal and interference levels. Hopping might take place in either time or frequency, the

only purpose being to randomize the interference levels, to improve code performance.

For such situations, it has been suggested that a simple way to make better use of the diversity antenna of a receiver is to adaptively select which antenna to use for transmission based on occasional measurements of signal strength [14]. Power control is then performed based on the antenna actually selected. Thus, selection diversity may be used in transmitting, with optimal combining used in reception. We assume that power control tracks shadowing but not Rayleigh fading.

When good estimates of signal strength and interference power can be made, soft decision decoding and antenna diversity combining using C/I estimates significantly outperform rival techniques. Simulations were run assuming perfect channel knowledge, and using transmit antenna selection diversity and receive antenna combining. Under these conditions, with the rate-1/4 code time hopping using techniques such as latin squares, all slots could be used in every cell, and still meet a 10^{-3} criterion for 1% outage. The outage probability followed a steep curve. By contrast, when the same interferers are encountered in each slot, the outage probability curve is fairly flat, already achieving 6.5% outage with full occupancy but declining to 1% outage only when half the slots were left vacant. Both results indicate higher capacity than in Example 1, reflecting the value of the perfect channel state information, and transmitter selection diversity. It is clear that here as well coding and interleaving when combined with latin squares significantly improves performance over using coding alone. Hopping in this example results in no variation in the level of the desired signal; rather, the benefits accrue from interference power level variations.

V. Conclusion

The propagation model outlined in Section II is too simple to be useful in predicting outage values in actual systems, and in any case we have made numerous idealizing assumptions (e.g., perfect power control). However, the model shares with all practical situations the feature that different interferers will have vastly different power levels. Thus, while the relative improvement due to use of schemes similar to the very simple method of latin squares in Section III may vary, we still expect dramatic improvement. The use of latin squares or similar methods to govern hopping in either frequency or time decreases the dependence of C/I across slots, allowing error correcting codes to realize more of their diversity benefit.

Systems which combine coding with latin squares are interference-limited rather than bandwidth-limited, in common with CDMA. That is, capacity is limited by the interference generated by adding more users, rather than in the number of slots allocated. This is true whether time-hopping or frequency-hopping or some combination thereof is employed. Thus, any technique which reduces interference or decreases the minimum C/I threshold for reliable operation directly increases the capacity (e.g., coding, antenna diversity, sectorization, voice activity). However, frequency hopping would be preferred in practice over time hopping since it affords less delay and provides the possibility of realizing frequency diversity without equalization.

We have assumed throughout that hops were synchronous in all cells. This is not actually required to achieve benefits from coding. However, the C/I estimates will not be as reliable if interference levels can change within a slot. Soft decision decoding would likely not be a realistic option unless the signal and interference conditions were very slowly changing and we could gather statistics over many successive frames.

Our results demonstrate that low-rate convolutional codes of very modest complexity may be used to achieve significant gain with low delay. A topic for further research is consideration of other codes. In particular, Reed Solomon codes may be attractive in situations where C/I estimates are poor, negating the soft decision advantage of convolutional codes. A recent study for frequency-hopped FSK indicates that convolutional codes are favored for the correct choice of metric [15], but further research with a variety of channel models is required to make a definite conclusion.

We are investigating one further possible use of latin squares in combination with coding. In this study, we have assumed that coding should be used to average the interference. However, time- or frequency-hopped systems may also be used to avoid interference. We are investigating distributed algorithms which dynamically select the M slots with best estimated C/I out of the N slots available in a frame. Thus, for the slots with the worst interference no signal is sent. Preliminary results indicate further improvement in capacity [16].

Acknowledgment

The authors thank J. E. Mazo and N. Seshadri of AT&T Bell Laboratories for their helpful comments on an earlier version of the manuscript. We also thank the anonymous reviewers for their helpful constructive comments.

References

[1] B. Gudmundson, J. Sköld, and J. K. Ugland, "A comparison of CDMA and TDMA systems," in *Proc. 1992 IEEE Veh. Technol. Conf.*, Denver, May 11–13, 1992, pp. 732–735.

[2] H. Sasaoka, "Block coded 16-QAM/TDMA cellular radio system using cyclical slow frequency hopping," in *Proc. 1992 IEEE Veh. Technol. Conf.*, pp. 405–408.

[3] K. S. Gilhousen *et al.*, "On the capacity of a cellular CDMA system," *IEEE Trans. Veh. Technol.*, vol. 40, pp. 303–312, May 1991.

[4] N. R. Livneh, R. Meridan, M. Ritz, and G. Silbershatz, "Frequency hopping CDMA for digital radio," in *Proc. Int. Commsphere Symp.*, Herzliya, Israel, Dec. 1991.

[5] D. R. Hughes and F. C. Piper, *Projective Planes*. New York: Springer-Verlag, 1973.

[6] A. D. Wyner, "Multi-tone multiple-access for cellular systems," AT&T Bell Laboratories Tech. Memo, BL011217-920811-11TM, Nov. 1991.

[7] J. Roche, "Families of binary matrices having nearly orthogonal rows," AT&T Bell Laboratories Tech. Memo, BL011217-920812-12TM.

[8] S. Lin and D. J. Costello, *Error Control Coding: Fundamentals and Applications*. Englewood Cliffs, NJ: Prentice-Hall, 1983.

[9] J. B. Anderson and S. Mohan, *Source and Channel Coding: An Algorithmic Approach*. Boston: Kluwer, 1991.

[10] J. W. Ketchum and M. S. Wallace, "Reed-Solomon coding and slow frequency hopping for narrowband digital cellular systems," in *Proc. 1992 IEEE Veh. Technol. Conf.*, pp. 225–228.

[11] N. Seshadri and C.-E. W. Sundberg, "Coded modulation with time diversity, unequal error protection, and low delay for the Rayleigh fading channel," presented at *Int. Conf. Universal Personal Commun.*, Dallas, Sept. 1992.

[12] A. R. Calderbank and N. Seshadri, "Multilevel codes for unequal error protection," *IEEE Trans. Inform. Theory*, vol. 39, no. 4, July 1993.

Performance Analysis of Antennas for Hand-Held Transceivers Using FDTD

M.A. Jensen and Y. Rahmat-Samii

IEEE Transactions on Antennas and Propagation
vol. 42, no. 8, pp. 1106-1113
August 1994

Performance Analysis of Antennas for Hand-Held Transceivers Using FDTD

Michael A. Jensen, *Student Member, IEEE*, and Yahya Rahmat-Samii, *Fellow, IEEE*

Abstract—**The design of antennas for hand-held communications devices depends on the implementation of simulation tools that can accurately model general topologies. This work presents the analysis of small antennas mounted on hand-held transceivers using the finite-difference time-domain (FDTD) method. The key features of the FDTD implementation are discussed, with particular emphasis placed upon modeling of the source region. The technique is used to predict the gain patterns and broadband input impedance behavior of monopole, planar inverted F, and loop antenna elements mounted on the handset. Effects of the conducting handset chassis, the plastic casing around the device, and lumped elements integrated into the antenna design are illustrated. Experimental results are provided to verify the accuracy of the computational methodology. The concept of antenna diversity is discussed, and key assumptions and expressions are provided that characterize the multipath fading fields. Several computational examples demonstrate the diversity performance of two receiving antennas on a single handset.**

I. Introduction

THE recent efforts aimed at improving available personal communications services have generated an increased interest in the performance of compact antenna structures mounted on hand-held devices. The characterization of such antennas is dependent upon the development of simulation tools that can accurately model general topologies, including wires, dielectrics, conductors, and lumped elements. In an effort to meet these simulation needs, attention has been focused on the use of the finite-difference time-domain (FDTD) [1], [2] methodology for antenna analysis [3]–[14]. Although past contributions in this arena have demonstrated the effectiveness of the FDTD approach in characterizing antenna configurations, only a limited amount of research has appeared relating to the simulation of practical antenna geometries operating in their true radiating environment.

This paper presents an investigation of the monopole, the planar inverted F antenna (PIFA) [15], and the loop antenna [16] (see Fig. 1) mounted on a hand-held transceiver using the FDTD technique. The radiation pattern, gain, and input impedance for these radiators are computed for several different topologies to illustrate the effects of the handset conducting chassis, the plastic casing surrounding the unit, and passive lumped elements integrated into the design on the antenna performance. A new source model is proposed that considerably reduces the time required to obtain the transient antenna response. Experimental data is provided for several of the configurations to validate the effectiveness of the FDTD for this class of problems. Due to page limitations, issues such as the interaction with the operator's tissue can not be detailed in this paper and as such are left to other publications [17].

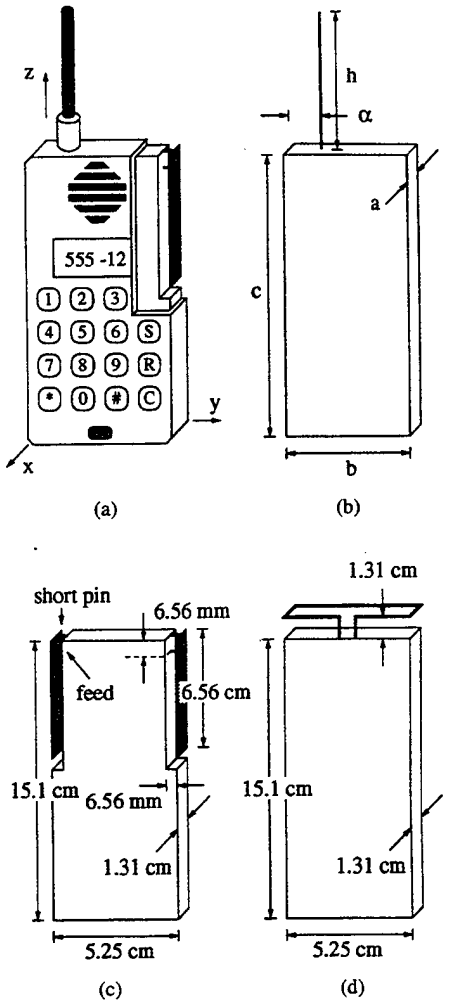


Fig. 1. (a) Typical handset geometry showing the monopole and PIFA elements. Configuration and dimensions for the (b) monopole, (c) dual PIFA, and (d) top-mounted loop on the handset.

Because hand-held communications equipment must often operate in a multipath fading environment, efforts to improve the data transfer reliability have motivated the implementation of antenna diversity schemes in these devices [18]–[24]. Therefore, the concept of antenna diversity is discussed, and key equations and assumptions used to describe the multipath

Manuscript received June 1, 1993; revised February 17, 1994. This work was supported by ARPA under Contract DAAB07-93-C-C501.

The authors are with the Electrical Engineering Department, University of California, Los Angeles, CA, 90024-1594 USA.

IEEE Log Number 9403658.

phenomenon are provided. Computational results are presented that illustrate the diversity performance of multiple small antennas mounted on a single handset.

II. FDTD Implementation

The FDTD [1] technique is a numerical approach that uses discrete approximations of Maxwell's time domain equations. The resulting algebraic equations can be used to track the time evolution of the fields within a given spatial region. The derivation as well as the practical implementation of this algorithm are well covered in past literature [1]–[14] and as such will not be covered in this paper.

In this work, the FDTD algorithm has been implemented to allow modeling and simulation of 1) very general radiating element geometries with dielectrics, wires, and planar conductors, 2) the conducting handset chassis, 3) the plastic casing surrounding the device, 4) the antenna feed, 5) lumped elements for tuning, and 6) multiple radiators for diversity. The technique is applied with Mur's second-order absorbing boundary condition at the outer computational domain boundary [11]. Far-zone gain and radiation patterns are computed using an integral near- to far-field transformation equation [12], [13]. The FDTD notation used in the following is similar to that provided in [9]. Note that Δx, Δy, and Δz denote the spatial grid discretization sizes and Δt represents the time step size.

A. Antenna Source Implementation

Typically, the excitation for wire-fed antenna geometries within the FDTD framework is performed using a gap voltage model in which a voltage is introduced in one cell of the feeding wire [3]. For a z-directed wire, the excitation electric field relates to the source voltage $V_s(t)$ through the expression

$$V_s(n\Delta t) = -E^n_{z,i,j,k}\Delta z \tag{1}$$

where the cell at (i, j, k) is located at the antenna feed point.

In this work, we present a new feed model that simulates a coaxial feeding cable. As seen in the side view of the coax illustrated in Fig. 2, a gap voltage is introduced in the coaxial center conductor, and the standard FDTD relations are used within the coax to propagate the fields toward the antenna. At the end of the coax, special interfacing relations are required to update the radial electric and circumferential magnetic fields. The relation for the circumferential magnetic field $H_{x,i,j,k}$ shown in Fig. 2 is derived using the integral forms of Maxwell's equations and assuming a $1/\rho$ dependence of the fields, as discussed in [8]. Using this approach, the modified time-stepping equation assumes the form

$$H^{n+1/2}_{x,i,j,k} = H^{n-1/2}_{x,i,j,k} + \frac{\Delta t}{\mu\Delta z}\left[E^n_{y,i,j,k} - E^n_{y,i,j,k-1}\frac{r_b}{\Delta y}\frac{\ln(r_b/r_a)}{\ln(\Delta y/r_a)}\right] - \frac{\Delta t}{\mu(\Delta y/2)\ln(\Delta y/r_a)}\left[E^n_{z,i,j,k} - E^n_{z,i,j-1,k}\right] \tag{2}$$

where r_a and r_b represent the radii of the coax inner and outer conductors, respectively. When computing the radial electric field $E_{y,i,j,k-1}$ in Fig. 2, the standard FDTD relations are used with the modification that $H_{x,i,j,k} \rightarrow (\Delta y/r_b)H_{x,i,j,k}$ to align $H_{x,i,j,k}$ with $H_{x,i,j,k-1}$. The parameters r_a, r_b, and ϵ within the coax are chosen to represent a desired characteristic impedance. Similar relations can be derived for the other circumferential magnetic and radial electric fields at the interface. Because reverse propagating modes in the coax are excited by the source voltage and by reflections from the coax-antenna transition, it is essential to terminate the coax with an absorbing boundary condition, as implied in Fig. 2.

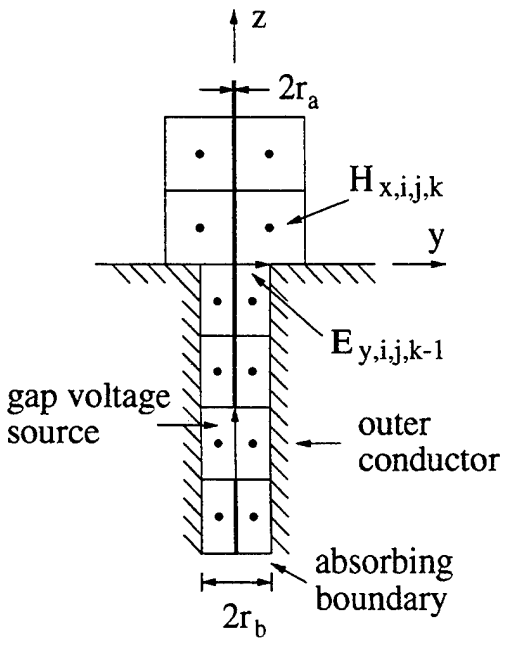


Fig. 2. Cross sectional view of the simulated coax and the grid used to determine the interface time-stepping equations.

For the simulated coax, the feed point voltage is obtained from the computed radial electric field using

$$V_s(n\Delta t) = \frac{r_b}{2}\ln\left(\frac{r_b}{r_a}\right)E^n_{y,i,j,k-1}. \tag{3}$$

The current in the wire for both source models can be obtained using a discretized form of Ampere's Law that assumes the form

$$I_s((n+1/2)\Delta t) = (H^{n+1/2}_{x,i,j,k} - H^{n+1/2}_{x,i,j+1,k})\Delta x + (H^{n+1/2}_{y,i+1,j,k} - H^{n+1/2}_{y,i,j,k})\Delta y. \tag{4}$$

The time offset of $\Delta t/2$ between the source voltage and current is generally neglected since it is assumed to be small.

Both source models discussed above appear to give virtually identical results in terms of input impedance and antenna gain. However, the simulated coaxial model provides two distinct advantages. First, it allows examination of the antenna input impedance at a point within the feeding transmission line rather than solely at the coax-antenna transition. Second, it has been observed in this study that the fields near the antenna terminals decay more rapidly in time for the coaxial model

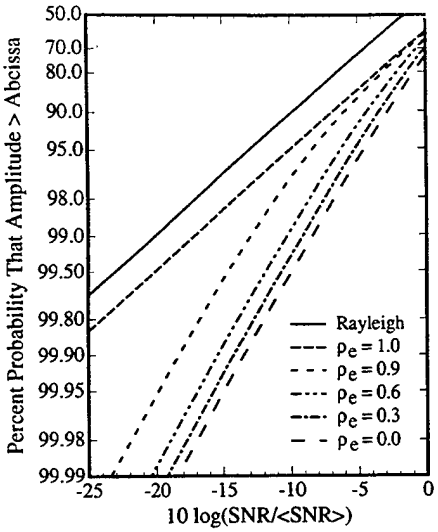


Fig. 3. Cumulative probability distribution for two-branch antenna diversity with maximal ratio combining for several values of the envelope correlation coefficient ρ_e. The Rayleigh reference for one antenna is provided for comparison.

than for the gap voltage model. In general, the time required for the terminal current to reach 1% of its peak value for the simulated coax is 30% to 60% of the time required for the standard gap model.

III. Antenna Diversity

The use of antenna diversity to mitigate the detrimental effects of signal fades in a multipath environment is a topic of considerable interest to designers in the personal wireless communications industry. For two antennas on a single receiver, the diversity performance is most commonly evaluated by investigating the correlation coefficient ρ_s–a statistical value indicating the similarity in the voltages received by the antennas. Because a great deal of literature exists that contains a statistical description of multipath fading fields as well as relations for the correlation coefficient [18]–[24], only a brief presentation relating to these topics will be presented here. In all expressions, the variables θ and ϕ are taken with respect to a coordinate system oriented with the z-axis perpendicular to the earth. The following basic assumptions concerning the incoming multipath fields are used in deriving expressions for the value of ρ_s:

1) The fading signal envelope is Rayleigh distributed.
2) Orthogonal polarizations are uncorrelated.
3) Each individual polarization is spatially uncorrelated.
4) The field arrives in the horizontal ($\theta = \pi/2$) plane only.
5) The time-averaged power density per steradian $\vec{S}(\theta,\phi) = S_\theta(\theta,\phi)\hat{\theta} + S_\phi(\theta,\phi)\hat{\phi}$ is constant in the horizontal plane such that $\vec{S}(\pi/2,\phi) = S_\theta^o\hat{\theta} + S_\phi^o\hat{\phi} =$ constant.

Using these assumptions, the correlation coefficient for the signals received by the two antennas can be expressed as [18]

$$\rho_s = \frac{\int_0^{2\pi} A_{12}(\phi)d\phi}{\left\{\int_0^{2\pi} A_{11}(\phi)d\phi \int_0^{2\pi} A_{22}(\phi)d\phi\right\}^{1/2}} \tag{5}$$

where

$$A_{pq}(\phi) = \Gamma E_{\theta p}(\pi/2,\phi)E_{\theta q}^*(\pi/2,\phi) + E_{\phi p}(\pi/2,\phi)E_{\phi q}^*(\pi/2,\phi) \tag{6}$$

and $\vec{E}_p(\theta,\phi) = E_{\theta p}(\theta,\phi)\hat{\theta} + E_{\phi p}(\theta,\phi)\hat{\phi}$ is the vector radiation pattern associated with antenna $\#p$. The envelope correlation coefficient, which is the measured quantity in antenna diversity experimentation, is approximately the magnitude squared of the voltage coefficient [18], or

$$\rho_e \approx |\rho_s|^2. \tag{7}$$

The parameter Γ in (6) is the cross-polarization discrimination (XPD) [21] of the incident multipath field and is given by

$$\Gamma = \frac{S_\theta^o}{S_\phi^o} \tag{8}$$

where S_θ^o and S_ϕ^o are identified in Assumption 5 above. In an urban environment, it is often assumed that either polarization is equally likely, resulting in a XPD of $\Gamma = 0$ dB [18]. However, there is still some speculation concerning the polarization characteristics of the incident field at the mobile antenna [22]–[24]. In light of this, the computations performed for this paper will be presented for a broad range of XPD values.

Fig. 3 shows the cumulative probability distribution of the signal-to-noise ratio (SNR) normalized to its time average ($\langle$SNR$\rangle$) for various values of ρ_e and for two branch antenna diversity. Maximal ratio combining, in which the two signals are received, co-phased, properly weighted, and added, is assumed in this plot. Also shown is the Rayleigh distribution that corresponds to a single antenna in the multipath environment. As can be seen, reducing the envelope correlation coefficient improves the probability distribution of the received SNR. For example, when no diversity is used, the Rayleigh distribution predicts that the received normalized SNR will be above -23 dB 99.5% of the time. If two antennas are used with $\rho_e = 0.6$, the SNR increases to -12 dB at the 99.5% reliability level. In light of this, the goal in diversity antenna design is to minimize ρ_e to the extent possible using a combination of spatial, angle, and polarization diversity.

The method for computing the vector patterns $\vec{E}_1$ and $\vec{E}_2$ depends upon the type of diversity combining used on the received signals. For example, if switched or selection diversity is to be implemented, the pattern $\vec{E}_1$ should be determined as the pattern for element #1 radiating in the presence of element #2 which is open circuited. For other combining techniques, the pattern should be computed for element #2 terminated with a matched load (absorbing boundary in the FDTD computation). For these computations, each pattern is obtained for the respective element radiating in the presence of a *parasitic* rather than *excited* second antenna in order to model the case of one antenna receiving while the second is positioned in a multipath fade [18].

IV. Computational and Experimental Results

In the following computations, the source function used is a sinusoid modulated by a Gaussian envelope that is expressed as

$$E(t) = \cos(\omega_o t)e^{-t^2/2\tau^2} \tag{9}$$

where the parameter τ controls the pulse width. Because this forcing function results in a frequency spectrum described by a Gaussian centered at ω_o and with variance $1/\tau^2$, it is possible to simply control the center frequency and bandwidth of the excitation. In the computations that follow, τ is chosen such that $1/\tau^2 =$ the bandwidth of interest. For single frequency computations, $\tau \rightarrow \infty$ is used.

The experimental measurements provided have been performed at the University of California, Los Angeles, antenna measurement laboratory. Impedance measurements are obtained from a Hewlett-Packard 8510B network analyzer. Pattern measurements are performed in a small anechoic chamber configured for far-field measurements. In all of the examples, the wire used has a radius r_a given by

$$2r_a = 0.9195 \text{ mm} \tag{10}$$

which corresponds to the radius of the inner conductor for RG402/U 50 Ω semirigid coaxial cable.

A. Monopole Antenna

The monopole is perhaps the most commonly used antenna for hand-held devices because of its broad-band characteristics and simple construction. Fig. 1(b) illustrates the geometry for such a configuration, where the monopole is centered on the handset in the x-direction and is offset a distance α in the y-direction. Fig. 4 compares the computed input impedance versus frequency with experimental measurements for the dimensions given in the figure caption. The small handset dimensions result from a frequency scaling that boosts the operating frequency of the device to 6 GHz. The computations performed use the simulated coaxial line discussed in Section II. As can be seen, very good agreement exists between the two sets of data, with only slight discrepancies occurring in the reactance near the upper end of the band. These curves also illustrate the broad-band nature of the monopole near its first resonance frequency, where the impedance is nearly matched to a 50 Ω feeding coax.

The effect of the conducting handset chassis on the monopole radiation pattern is an important consideration in the design of antennas for practical applications. Fig. 5 provides the $\hat{\theta}$-polarized patterns normalized to the antenna gain for the monopole on the handset at a frequency of 6 GHz. The experimental data shown extends only across the upper hemisphere because brackets in the measurement facility interfered with the lower hemisphere measurements. Here again, good correlation exists between the computed and measured results. For comparison purposes, the gain pattern for a 2.5-cm dipole antenna, whose pattern shape corresponds to that of the monopole on an infinite ground plane, is also shown. As can be seen, the presence of the chassis alters the radiation pattern. Simulations of this nature

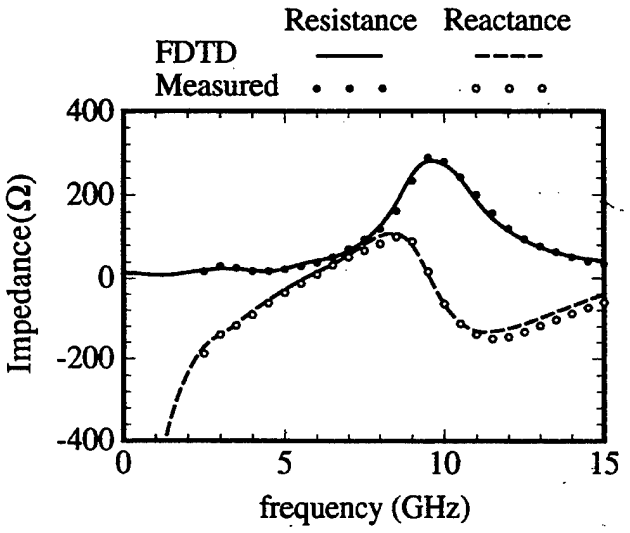


Fig. 4. Computed and measured input impedance versus frequency for the monopole on the handset illustrated in Fig. 1(b) for the parameters $a = 1$ cm, $b = 2$ cm, $c = 3$ cm, $h = 1.25$ cm, and $\alpha = 1$ cm. These small dimensions result from frequency scaling the handset to operate at 6 GHz.

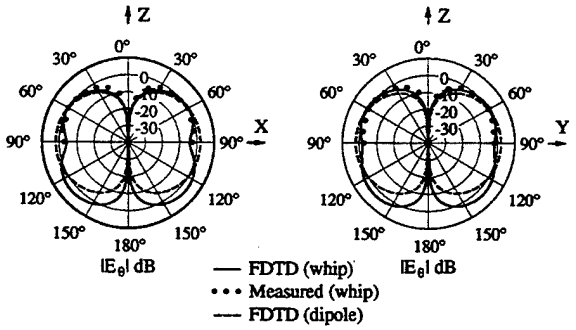


Fig. 5. Computed gain patterns at 6 GHz compared to measured data in the upper hemisphere for the geometry of Fig. 1(b) with the dimensions given in the caption of Fig. 4. The gain patterns for a 2.5-cm dipole are provided for comparison.

provide considerable insight into the radiation behavior of the antenna/transceiver system.

In many applications, an inductive load is integrated into the monopole to improve the match or reduce the antenna size. Fig. 6 illustrates this effect by presenting the variation of $|S_{11}|$ (assuming a 50 Ω feeding transmission line) with frequency for the monopole of Fig. 1(b). The handset dimensions are again provided in the figure caption. The different curves represent the results when the monopole is alone on the handset, when a plastic casing is present, and when the plastic and a 10-nH inductor one FDTD cell above the monopole feed point are included. The plastic layer is modeled as a 3.28 mm thick lossless dielectric with permittivity $\epsilon_r = 2$, which is immediately adjacent to the conducting chassis. An expanded view of the plot about the 900-MHz resonance is provided for clarity. These results show that the plastic exercises relatively little effect on the monopole antenna performance at the lower end of the frequency band and a slightly more significant effect at higher frequencies. Also, the inductive load reduces the resonant frequencies and improves the match, especially at

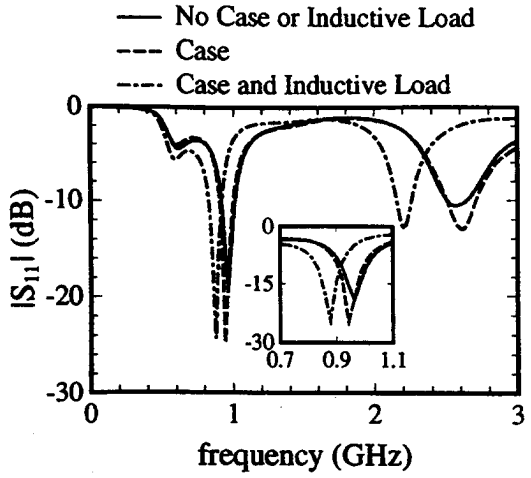


Fig. 6. Effect of the 3.28-mm plastic case and 10-nH inductor on the variation of $|S_{11}|$ versus frequency for the monopole of Fig. 1(b). The dimensions are $a = 1.31$ cm, $b = 5.25$ cm, $c = 15.1$ cm, $h = 8.5$ cm, and $\alpha = 6.56$ mm.

the higher frequencies. This capability of modeling lumped elements within the FDTD simulation tool allows the design engineer to efficiently determine proper loading strategies to improve the performance of antennas for communications devices.

B. Planar Inverted F Antenna

An interesting antenna topology that has received some attention in the literature is the planar inverted F antenna (PIFA), a modified microstrip that allows a simple impedance match in a low-profile design [15], [25], [26]. This geometry is illustrated in Fig. 1(c), where the antenna is mounted on the side of a conducting handset chassis. The short circuit introduced at one edge of the radiating element allows the antenna size to be reduced by a factor of two for a given resonant frequency. Past observations have revealed that the PIFA offers bandwidths as high as 10% when it is located on a small ground plane or conducting chassis [15].

Fig. 7(a) depicts the geometry for a PIFA element on a finite ground plane. The input impedance variation with frequency for this configuration is illustrated in Fig. 7(b). Measured data obtained from a prototype of this antenna configuration is also presented in the figure. As can be observed, very good agreement exists between the two sets of data. This plot also demonstrates the excellent match between the antenna and coax near 900 MHz.

For an integrated antenna such as the PIFA, the effect of the plastic case that encloses the handset is an important consideration. Fig. 8 illustrates this by presenting the input impedance versus frequency for the PIFA of Fig. 1(c) with and without the plastic casing. Referring to Fig. 1(c), the antenna located toward the positive y-axis is excited while the second is terminated with an absorbing boundary. The plastic layer is again a 3.28 mm thick lossless dielectric with permittivity $\epsilon_r = 2$. These results reveal that the dielectric loading by the casing noticeably reduces the antenna resonant frequency and therefore allows the element size to be reduced for a given

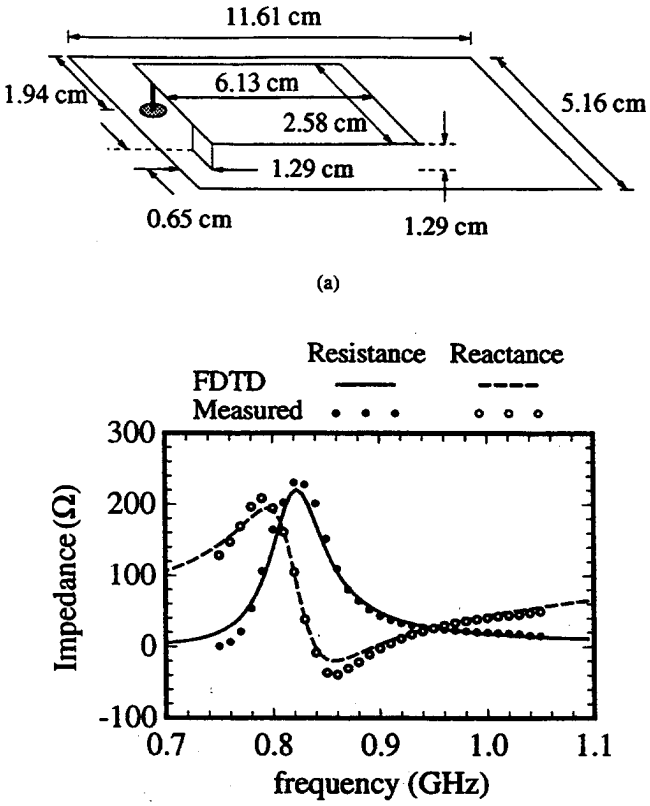


Fig. 7. PIFA on a finite ground plane: (a) geometry and (b) computed and measured input impedance versus frequency.

operating frequency. Fig. 9 represents the gain patterns at 915 MHz in the principal planes for the PIFA element of Fig. 1(c) with the plastic casing. As can be seen, the PIFA provides a pattern very similar to that of the monopole.

C. Loop Antenna

The loop antenna is an element that has often been used in pager devices, but to date it has found very little application in hand-held transceivers. This is probably because the small loop is highly inductive and difficult to match to the feed line. However, as the operating frequency of wireless communications devices moves into higher bands, the loop becomes a viable antenna element for these applications, particularly in designs where balanced amplifiers must interface with the antenna. Fig. 1(d) illustrates the geometry for a wire loop mounted on the top of a handset. Fig. 10 presents the input impedance variation for this configuration both with and without the 3.28 mm thick plastic casing, which completely encloses the handset and antenna. Once again, the dielectric lowers the resonance frequencies of the radiator. This plot further shows that though the loop presents challenging matching requirements at low frequencies, its impedance characteristics are well behaved near 2 GHz, where the loop perimeter approaches 1λ.

D. Diversity Performance of Dual Antennas

Determining the diversity performance of two antennas mounted on a handset requires computation of the envelope

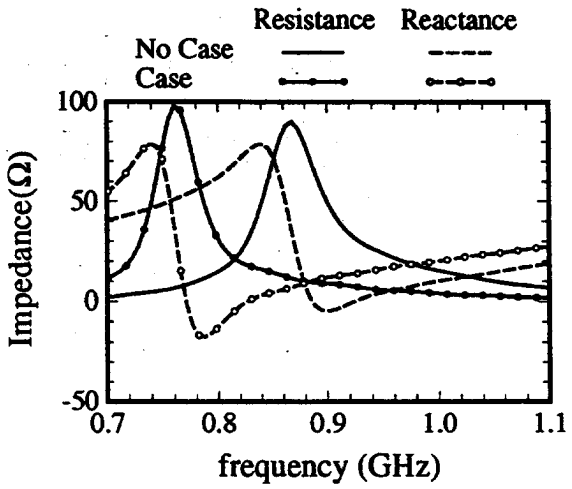


Fig. 8. Input impedance versus frequency for the side-mounted PIFA in Fig. 1(c) with and without a 3.28-mm plastic casing enclosing the handset.

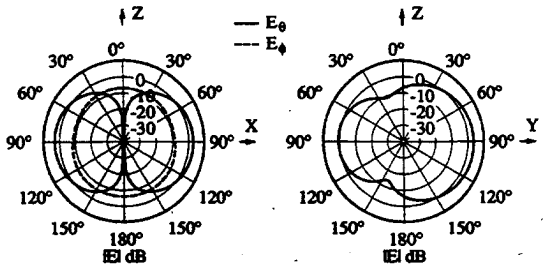


Fig. 9. Gain patterns at 915 MHz for the side-mounted PIFA of Fig. 1(c) with the 3.28-mm plastic casing.

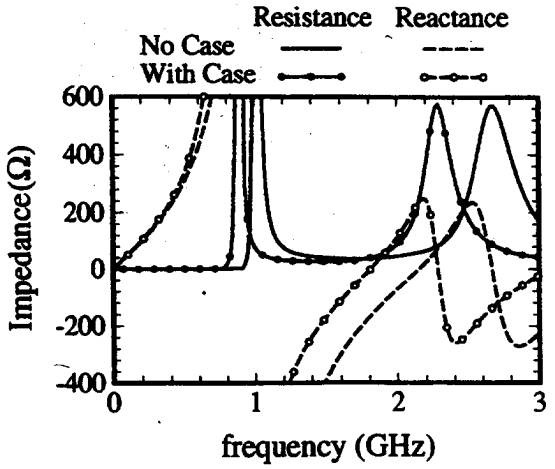


Fig. 10. Input impedance versus frequency for the loop on the handset illustrated in Fig. 1(d) with and without the 3.28-mm plastic casing.

correlation coefficient. Consistent with assumption 4 in Section III, it is assumed that the incident multipath field is confined to a plane parallel with the horizon. To illustrate the effects of varied orientation during operation, the handset coordinate system is allowed to rotate about its x-axis such that its z-axis makes an angle β with the perpendicular to the horizon. As a first example of this computation, the dual PIFA geometry of

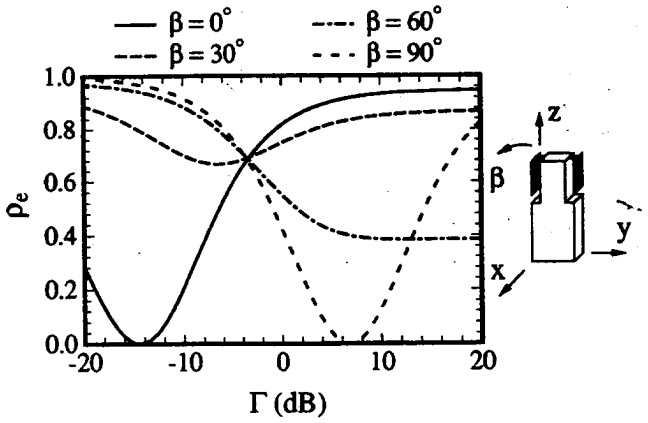


Fig. 11. Envelope correlation coefficient versus the XPD (Γ) at 915 MHz for the dual PIFA configuration shown in Fig. 1(c) for several different handset orientations.

Fig. 1(c) is used with the plastic casing included. The radiation patterns at 915 MHz computed for each antenna (with the other antenna terminated in a matched load) are incorporated into (5) and (7) to obtain ρ_e. The resulting values of ρ_e are illustrated versus the XPD (Γ) of the incident multipath field in Fig. 11 for several values of β. The dependence of ρ_e on Γ implies that significant decorrelation is produced using polarization diversity between the two branches. Also noteworthy is the influence of pattern diversity as the handset is rotated.

A second interesting example is the combination of the monopole in Fig. 1(b) with one of the PIFA elements in Fig. 1(c), as implied by Fig. 1(a). The monopole dimensions are those used in the caption of Fig. 6, and the plastic casing is again included. The results are shown in Fig. 12 at 915 MHz versus Γ for several different handset orientations. Note that in this case the value of ρ_e is lower for $\Gamma = 0$ dB than for the case of the dual PIFA antenna studied in Fig. 11. This improvement is attributed to the fact that, in contrast to the dual PIFA configuration, the two elements used for Fig. 12 exhibit different pattern and polarization characteristics.

As a demonstration of the use of Figs. 11 and 12 for determining the benefits of the diversity schemes, consider the curve for $\beta = 30°$ in Fig. 12, which predicts that $\rho_e \approx 0.3$ at $\Gamma = 0$ dB. Using this value in Fig. 3 suggests that the received SNR for the diversity configuration is approximately 12 dB higher than that for a single antenna in a Rayleigh fading environment. Studies such as this aid the antenna designer in selecting antenna configurations that will be effective in combating rapid fading due to multipath delays in urban environments.

V. Conclusion

In this work we have used the FDTD method to accurately characterize the performance of single- and dual-antenna configurations mounted on hand-held communications devices. The key features of the FDTD implementation have been presented, with special consideration given to the source models used. The concept of antenna diversity has also been

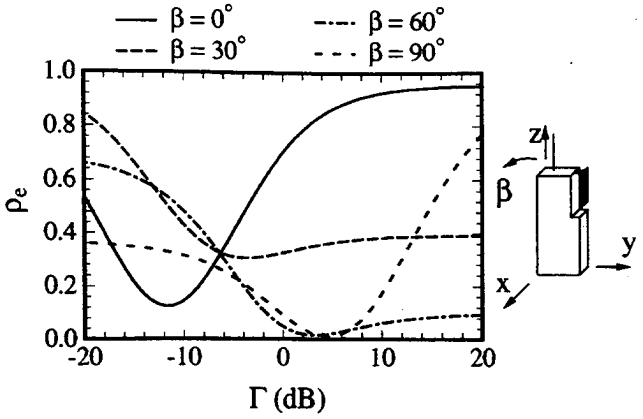


Fig. 12. Envelope correlation coefficient versus the XPD (Γ) at 915 MHz for the PIFA-monopole configuration shown in Fig. 1(a) for several different handset orientations.

discussed, and an expression relating the envelope correlation coefficient to the radiation patterns of two coupled antennas has been provided. The geometrical flexibility of the FDTD algorithm has allowed simulation of the monopole, PIFA, and loop elements operating in the presence of a conducting chassis and plastic casing similar to what might form the body of a transceiver handset. Comparisons have demonstrated that the FDTD results match very well with measured data for both radiation patterns and broad-band input impedance. We have also demonstrated the use of the FDTD tool in determining the diversity performance of two antennas operating simultaneously on the handset.

ACKNOWLEDGMENT

The authors thank J. Colburn for his help in obtaining the experimental data.

REFERENCES

[1] K. S. Yee, "Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media," *IEEE Trans. Antennas Propagat.*, vol. AP-14, pp. 302–307, May 1966.

[2] A. Taflove and M. E. Brodwin, "Numerical solution of steady-state electromagnetic scattering problems using the time-dependent Maxwell's equations," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-23, pp. 623–630, Aug. 1975.

[3] R. Luebbers, L. Chen, T. Uno, and S. Adachi, "FDTD calculation of radiation patterns, impedance, and gain for a monopole antenna on a conducting box," *IEEE Trans. Antennas Propagat.*, vol. 40, pp. 1577–1583, Dec. 1992.

[4] J. G. Maloney, G. S. Smith, and W. R. Scott, Jr., "Accurate computation of the radiation from simple antennas using the finite-difference time-domain method," *IEEE Trans. Antennas Propagat.*, vol. 38, pp. 1059–1068, July 1990.

[5] A. Reineix and B. Jecko, "Analysis of microstrip patch antennas using finite difference time domain method," *IEEE Trans. Antennas Propagat.*, vol. 37, pp. 1361–1369, Nov. 1989.

[6] D. S. Katz, M. J. Piket-May, A. Taflove, and K. R. Umashankar, "FDTD analysis of electromagnetic wave radiation from systems containing horn antennas," *IEEE, Trans. Antennas Propagat.*, vol. 39, pp. 1203–1212, Aug. 1991.

[7] P. A. Tirkas and C. A. Balanis, "Finite-difference time-domain method for antenna radiation," *IEEE Trans. Antennas Propagat.*, vol. 40, pp. 334–340, Mar. 1992.

[8] A. Taflove, K. R. Umashankar, B. Beker, F. Harfoush, and K. S. Yee, "Detailed FD-TD analysis of electromagnetic fields penetrating narrow slots and lapped joints in thick conducting screens," *IEEE Trans. Antennas Propagat.*, vol. 36, pp. 247–257, Feb. 1988.

[9] D. M. Sheen, S. M. Ali, M. D. Abouzahra, and J. A. Kong, "Application of the three-dimensional finite-difference time-domain method to the analysis of planar microstrip circuits," *IEEE Trans. Microwave Theory Tech.*, vol. 38, pp. 849–857, July 1990.

[10] W. Sui, D. A. Christensen, and C. H. Durney, "Extending the two-dimensional FDTD method to hybrid electromagnetic systems with active and passive lumped elements," *IEEE Trans. Microwave Theory Tech.*, vol. 40, pp. 724–730, Apr. 1992.

[11] G. Mur, "Absorbing boundary conditions for the finite-difference approximation of the time-domain electromagnetic field equations," *IEEE Trans. Electromagn. Compat.*, vol. EMC-23, pp. 377–382, Nov. 1981.

[12] R. J. Luebbers, K. S. Kunz, M. Schneider, and F. Hunsberger, "A finite-difference time-domain near zone to far zone transformation," *IEEE Trans. Antennas Propagat.*, vol. 39, pp. 429–433, Apr. 1991.

[13] R. J. Luebbers and J. Beggs, "FDTD calculation of wide-band antenna gain and efficiency," *IEEE Trans. Antennas Propagat.*, vol. 40, pp. 1403–1407, Nov. 1992.

[14] M. A. Jensen and Y. Rahmat-Samii, "FDTD analysis of PIFA diversity antennas on a hand-held transceiver unit," in *1993 IEEE AP-S Int. Symp. Dig.*, Ann Arbor, MI, June 27–July 2, 1993, vol. 2, pp. 814–817.

[15] T. Taga, "Analysis of planar inverted-*F* antennas and antenna design for portable radio equipment," in *Analysis, Design, and Measurement of Small and Low Profile Antennas*, K. Hirasawa and M. Haneishi, Eds. Boston: Artech, 1992, ch. 5.

[16] M. A. Jensen and Y. Rahmat-Samii, "Characterisation of electromagnetically coupled superquadric loop antennas for mobile communications applications," *IEE Proc. H*, vol. 141, pp. 85–93, Apr. 1994.

[17] ——, "The electromagnetic interaction between biological tissue and antennas on a transceiver handset," in *1994 IEEE AP-S Int. Symp. Dig.* Seattle, WA, June 19–24, 1994, pp. 367–370.

[18] R. G. Vaughan and J. B. Andersen, "Antenna diversity in mobile communications," *IEEE Trans. Veh. Technol.*, vol. VT-36, pp. 149–172, Nov. 1987.

[19] W. C. Jakes, Jr., *Microwave Mobile Communications*. New York: Wiley, 1974.

[20] R. G. Vaughan, "Polarization diversity in mobile communications," *IEEE Trans. Veh. Technol.*, vol. 39, pp. 177–186, Aug. 1990.

[21] P. C. F. Eggers, J. Toftgard and A. M. Oprea, "Antenna systems for base station diversity in urban small and micro cells," *IEEE J. Select. Areas Commun.*, vol. 11, pp. 1046–1057, Sept. 1993.

[22] S. Kozono, T. Tsuruhara, and M. Sakamoto, "Base station polarization diversity reception for mobile radio," *IEEE Trans. Veh. Technol.*, vol. VT-33, pp. 301–306, Nov. 1984.

[23] W. C. Y. Lee and Y. S. Yeh, "Polarization diversity system for mobile radio," *IEEE Trans. Commun.*, vol. COM-20, pp. 912–922, Oct. 1972.

[24] R. G. Vaughan, "Signals in mobile communications: A review," *IEEE Trans. Veh. Technol.*, vol. VT-35, pp. 133–145, Nov. 1986.

[25] T. Taga and K. Tsunekawa, "Performance analysis of a built-in planar inverted *F* antenna for 800 MHz band portable radio units," *IEEE J. Selected Areas Commun.*, vol. SAC-35, pp. 921–929, June 1987.

[26] K. Kagoshima, K. Tsunekawa, and A. Ando, "Analysis of a planar inverted *F* antenna fed by electromagnetic coupling," in *1992 IEEE AP-S Int. Symp. Dig.*, Chicago, July 18–25, 1992, pp. 1702–1705.

Michael A. Jensen (S'92) received the B.S. (*summa cum laude*) and M.S. degrees in electrical engineering from Brigham Young University (BYU), Provo, UT, in 1990 and 1991, respectively. He is currently pursuing the Ph.D. degree in electrical engineering at the University of California, Los Angeles (UCLA).

From 1989 to 1991 he was a Graduate Research Assistant in the Lasers and Optics Laboratory, BYU. In 1990 he received a National Science Foundation Graduate Fellowship. Since 1991, he has been with the Antenna Laboratory at UCLA as a Graduate Student Researcher. His main research interests include optical fiber communication, radiation and propagation for personal communications, numerical electromagnetics, and implementation of finite-difference schemes on massively parallel computer architectures.

Mr. Jensen is a member of Eta Kappa Nu and Tau Beta Pi.

Yahya Rahmat-Samii (S'73-M'75-SM'79-F'85) received the M.S. and Ph.D. degrees in electrical engineering from the University of Illinois, Champaign-Urbana.

He is a Professor of Electrical Engineering at the University of California, Los Angeles. He has been a Senior Research Scientist at NASA's Jet Propulsion Laboratory/California Institute of Technology since 1978, where he contributed significantly to the advancement of antenna technology for space programs. He was a Guest Professor at the Technical University of Denmark (TUD) in the summer of 1986. He has also been a consultant to many aerospace companies.

Dr. Rahmat-Samii is a Fellow of IAE (1986) and was the 1984 recipient of the Henry Booker Award of URSI. He was appointed an IEEE Antennas and Propagation Society Distinguished Lecturer and presented lectures internationally. He was an elected IEEE AP-S AdCom member for the second term and has been an Associate Editor of the IEEE TRANSACTIONS ON ANTENNAS AND PROPAGATION and the society's magazine. He is currently the elected Vice President of IEEE AP-S. He was the Chairman of the IEEE Antennas and Propagation Society of Los Angeles in 1987–1989. In 1989, his chapter received the Antennas and Propagation Best Chapter Award from the AP Society. He is one of the three International Editors of the IEE book series on Electromagnetics and Antennas. He is also one of the Editors of the *Journal of Electromagnetic Waves and Applications*. He is one of the Directors of the Antenna Measurements Technique Association (AMTA) and the Electromagnetics Society. He is listed in *Who's Who in America*, *Who's Who in Frontiers of Science and Technology*, and *Who's Who in Engineering*. He has authored or coauthored over 300 technical journal articles and conference papers and has written chapters in 12 books. He has made pioneering contributions to the developments of near-field plane-polar and bipolar antenna measurements, microwave holographic diagnostics, mobile satellite communication antennas, reflector surface compensation, multireflector antenna diffraction analysis and synthesis, scattering and radiation from complex objects, RCS computations, singularity in dyadic Green's function, high-power microwave (HPM) antennas, EMP and aperture penetration, the Spectral Theory of Diffraction (STD), and GTD. For these contributions he has received numerous NASA Certificates of Recognitios and recently earned the JPL Team NASA's Distinguished Group Achievement Award. In 1992, he was the recipient of the Best Application Paper Award (Wheeler Award) for a paper published in 1991. In 1993, one of his Ph.D. students was named the Most Outstanding Ph.D. Student at UCLA's School of Engineering and Applied Science and another received the Best Student Paper Award at the 1993 IEEE AP-S Symposium. He is a member of Commissions A, B, and J of USNC/URSI, Sigma Xi, Eta Kappa Nu, and the Electromagnetics Academy.

Low Power RF-ICs in Wireless Transceivers

A.A. Abidi

Proceedings 1994 IEEE Symposium on Low Power Electronics
pp. 18-21
August 1994

Low-Power RF-ICs in Wireless Transceivers

Asad A. Abidi
Electrical Engineering Department
University of California
Los Angeles, CA 90024-1594

Introduction

It is anticipated that about 300 million portable consumer wireless devices will be in use by the year 2000 [1]. What IC technologies will enable the RF front-end of these devices? Do miniaturization and long battery-life call for architectural innovations in transceivers? What new circuit design styles will evolve in response?

There is much curiosity and speculation on these matters, yet little is known in general terms about RF-IC design, and the impact of large-scale integration and power-reduction strategies in front-ends of wireless transceivers. This paper presents a brief survey of the highlights in the application of IC technology to radios since the 1970s, and follows this with trends of what the author considers is the shape of things to come in wireless transceivers over the next few years. RF-ICs are roughly defined as integrated circuits operating in the band of frequencies from 100 MHz to 2000 MHz which covers most consumer wireless devices. MMICs, by contrast, are small-scale ICs operating at tens of GHz which are almost exclusively fabricated on III-V compound semiconductor substrates. Driven by defence applications over the past decade, they have become an established discipline. RF-ICs are principally driven by consumer needs. They are, and will continue to be, fabricated mostly in mature silicon technologies, combining analog and digital functions at relatively large levels of integration. It is the authors belief that in response to pressing demands for ubiquitous wireless access, the day of the "VLSI radio" is not that far off.

ICs in Broadcast Radio Receivers

Integrated circuits are used in one form or another in the front-end of most of today's broadcast receivers. However, it is fair to say that aside from contributing building blocks as the double-balanced mixr and the phase-locked loop, IC technology has *not* caused a profound re-thinking of radio receivers, whose architecture still resembles post-World War II radios. Some recent advances in this respect are now described.

The two most important specifications of a receiver are its *selectivity* and its *sensitivity*. Selectivity is the receiver's ability to reject adjacent channels and image frequencies, while sensitivity is its ability to detect weak input signals. The noise properties of the transistors in the front-end mainly set the sensitivity, whereas the receiver's frequency-allocation plan and quality of filters determine its selectivity. The frequency-allocation plan consists of the choice of the intermediate frequencies (IF) and of filtering to obtain a desired image suppression. In an FM receiver, for instance, as it is difficult to make high selectivity filters at RF, a first IF of 10.7 MHz is used to ensure that the unwanted image that will superpose itself on the desired channel lies outside the FM band of 88-108 MHz, and is inherently rejected by the detector. Passive crystal filters with a 150 kHz passband centered at 10.7 MHz then reject adjacent channels. In designing a low-power, miniature FM receiver, Kasperkovitz [2] at Philips realized there are many advantages in eliminating the RF and IF filters, and implementing the required selectivity on-chip. As it was impossible to integrate the RF filters on-chip with spiral monolithic inductors, or to replace the 10.7 MHz IF crystal filters with active gyrator filters, the key step in solving this problem was to *lower the IF* from 10.7 MHz to 70 kHz. Aside from a small off-chip inductor which provides a minimum of RF filtering, the channel bandwidth is mainly defined on-chip by high-order active *RC* filters. The power dissipation in these filters for a given dynamic range scales down with the IF [3], as does the required operating current in the IF amplifier. The image frequency, instead of lying outside the FM band, now lies *between* FM channels (Fig.1), so it only superposes noise on to the desired signal. However, a consequence of this choice is that the ±75 kHz frequency deviation of the desired FM channel may alias around DC to create distortion, so it is compressed to ±15 kHz by a negative frequency feedback. This single-chip FM mono receiver [4] requires only 15 small capacitors and two inductors, making it possible to fit the entire system within a wristwatch. The IC drains 8 mA from a 4.5V battery supply.

The same idea of low-IF has been pursued at Sony for their personal entertainment broadcast receivers. The bipolar RF-ICs operate at single-battery supplies as low as 0.95 V. A recent FM radio IC implements an unusual double-superheterodyne architecture [5]. At 30 MHz, the first IF is high enough for image-suppression by a fixed, low-selectivity bandpass filter after the antenna spanning the entire broadcast range from 88 to 108 MHz; no on-chip filter is used. Then, after amplification at the 30 MHz IF which comes readily in a state-of-the-art bipolar technology, a second *image-rejection mixer* (Fig.2) converts to a low IF of 150 kHz, and thereafter on-chip filtering is implemented by a 9th order active *RC* lowpass section. The RF amplifier transistors are integrated on-chip, although they require off-chip inductors for the load, and the local oscillator needs an off-chip *LC* tuned circuit. This receiver drains about 15 mA in either FM or AM mode.

ICs in Wireless Paging Receivers

Paging receivers have been in continuous development since the early 1960's, and have evolved into today's wide-area pagers which use digital signalling. The modern pager signals anywhere from 500 to 1200 b/s by encoding data with a simple positive or negative offset of the carrier frequency — the binary frequency-shift keyed (FSK) modulation. Paging receivers up to the early 1980's did not

The UCLA Low-Power Wireless Transceiver Project is supported by the US Advanced Research Projects Agency, and by a consortium of semiconductor companies under the State of California MICRO Program.

0-7803-1953-2 / 94 / $3.00 / © 1994 IEEE

exploit the powerful simplifications this signaling scheme implied. Vance at Standard Telecommunication Laboratories realized how to simplify an FM receiver on to one chip with a quadrature downconversion mixer combined with *zero-IF*. Although this might not be suitable for a broadcast FM receiver, it was exactly the right solution for the binary-FSK digital paging receiver [6]. The remarkably simple receiver (Fig.3) consists only of a quadrature demodulator, lowpass filters in each arm, limiters, and a D-type flip-flop detector, and when integrated on one bipolar chip, it drains 2.5 mA from 1.8 V when active; in standby the drain falls to a mere 50 μA. More sophisticated detectors may be used to gain better noise rejection. Zero-IF means that the the image is entirely eliminated. DC offsets in the receiver or flicker noise now fall in-band, however, but as the data is encoded in frequencies ±4.5 kHz away from the carrier, the receiver may be AC coupled. The 10 kHz lowpass filters serve to eliminate out-of-band noise and blocking signals which would degrade the signal-to-noise ratio at the detector.

Single-chip bipolar zero-IF paging receivers are now quite common. A prominent feature they all share is that much of the die area is taken up by the capacitors for the on-chip lowpass filters and AC coupling [7, 8]. When active, these receivers consume only 1 to 2 mA from a single-battery, which is the typical specification for pagers. This example shows how the zero-IF, direct conversion architecture proved to be a good match to pager signalling.

ICs in Cellular Telephone Transceivers

Mobile and handheld cellular telephones are the first widespread two-way radios for consumer use, for which cordless telephones were a prelude. To be acceptable to the consumer, radio telephones must meet stringent demands for low weight and volume, long battery life, and low-cost. Further, to support a large numbers of users in a crowded radio spectrum, radio telephones require more signal processing than do other common transceivers. They must therefore use high levels of integration in the electronics.

The first generation of ICs for use in cellular telephones contained the frequently occurring building blocks used in conventional receivers, such as the mixer and local oscillator, or the IF amplifier chain and signal-strength indicator. Analog cellular telephones typically use the double-conversion architecture (Fig.6), assembled with these building blocks, often using GaAs ICs at the receiver and transmitter modules, silicon building blocks in the IF, surrounded by an impressive collection of discrete resonator filters and antenna duplexers [9].

Circuits to implement the complex modulation formats in the new generation of digital mobile telephones, with their greater requirements to withstand nearby blocking signals, require yet greater integration of the RF and IF electronics. In a 900 MHz transmitter silicon bipolar IC from Siemens [10], a precision quadrature-upconversion mixer implements the phase-shift keyed modulation specified in the European GSM system to produce a single-sideband, suppressed-carrier output. The IC, which includes an output stage to drive a separate power amplifier, drains 40 mA. The companion receiver IC includes a first mixer, with buffers to drive an IF off-chip bandpass SAW filter, and a quadrature downconversion mixer to detect the single-sideband input at baseband. Including an AGC with 70 dB range operating at the 40 to 90 MHz IF and a signal-strength indicator, the receiver drains 24 mA. Although the on-chip gain may become as large as 80 dB, it is distributed in different frequency bands and as a result there is little on-chip crosstalk.

Yet greater miniaturization is attained at Alcatel by integrating the transmitter and receiver sections together on one silicon bipolar chip. The zero-IF architecture makes this possible [11]. An off-chip low-noise amplifier drives directly into quadrature (I-Q) mixers, whose dynamic range is wide enough to withstand large blocking signals 3 MHz away from the carrier without producing significant intermodulation. Large (650 pF) off-chip capacitors constitute an anti-alias filter. The downconverted signals are sent to a baseband CMOS IC, containing a high-order switched capacitor lowpass filter to remove the blocking signals, and a demodulator. An *RC* and *CR* network with off-chip trimming shifts the local oscillator phase by ±45° to produce the quadrature drive to the mixers. Separate upconversion mixers in the transmitter drive a power amplifier module. The chip drains 25 mA from 5V in the receive mode, and 45 mA in the transmit mode, rather comparable to the Siemens chip set. It also suppresses the image and carrier by about 40 dB.

ICs for Spread-Spectrum Wireless Transceivers

As cellular telephones saturate the available radio frequency allocations, spread-spectrum techniques are being deployed in cordless telephones and wireless modems. These usually operate in the two lower bands allocated by the FCC for the industrial, scientific, and medical (ISM) communities for unlicensed spread-spectrum communications: 902 to 928 MHz, and 2.4 to 2.83 GHz. A ubiquitous wireless environment is envisioned, in which mobile users, wherever they are, may access data and communications services through an intricate network of base stations. The greatest hardware challenge is in developing the low-power, miniature handset.

A notable recent example of a miniature spread-spectrum transceiver is the 700 kb/s frequency-hopped device from Plessey, which operates in the 2.4 GHz band, and is assembled on 2"×3" PCMCIA card for insertion into notebook computers [12]. It communicates by binary frequency-shift keying of the carrier, and to spread the spectrum, the carrier frequency, in turn, is slowly hopped across the entire band by a variable-modulus PLL synthesizer. All the active devices in the transceiver are on *three ICs*, consisting of a GaAs RF front-end, a silicon bipolar IF receiver, and a CMOS IC for the hopping frequency-synthesis. Nevertheless, the transceiver requires 50 passive components, including six rather bulky passive components such as filters, and it dissipates more than 1 W while transmitting 100 mW.

The single-chip GaAs IC front-end in this transceiver [13] is one of the most highly integrated, as it includes the power amplifier and drivers for 2-GHz passive filters in the transmit and receive paths. Otherwise, it is a conventional double-superheterodyne architecture. In its first version, the IC has more than 20 on-chip spiral inductors, probably the largest number found on any MMIC! In receive mode, the IC drains 30 mA from a 5 V supply, and it selects one of two antennas with an on-chip RF switch for *spatial diversity*.

ICs to Enable Future Transceivers

What might the portable wireless communicator of the future look like? We may draw some conclusions from the foregoing summary of RF-IC developments. The crowded spectrum means that wireless communicators will predominantly use spread-spectrum techniques, and that the need for low power dissipation will force yet higher levels of integration. Much of the filtering will be done on-chip at low frequencies, with perhaps only one RF passive filter off-chip. The author with his colleagues and their graduate students at UCLA is investigating the architecture and circuit design of a frequency-hopped, binary frequency-shift keyed, zero IF, all-CMOS two-chip transceiver capable of delivering up to 160 kb/s (the base ISDN rate) in the 900 MHz ISM band (Fig.5) [14]. This approach is to a large degree inspired by the modern paging receiver, which is perhaps the lowest-energy wireless device in wide use. The transceiver ICs freely mix analog and digital circuits, making CMOS the IC technology of choice for the entire transceiver, *including* the RF front-end. These blocks will be implemented in an unmodified standard production 1-μm CMOS process.

Low-power operation requires that the entire system should operate on a 3V supply, which cannot be any lower because of the analog sections, and the unmodified FET threshold voltages. The power otherwise wasted in driving pad and trace parasitic capacitances is avoided with the use of on-chip inductors, whose self-resonance after the process of selective removal of the silicon substrate (Fig.6) [15] is as high as it is on semi-insulating GaAs substrates. Thus, a 1-μm CMOS 900 MHz RF amplifier (Fig.6) has 30 dB gain and drains only 3 mA. The input FETs are biased at low V_{GS}–V_t to attain a large g_m/I_D. The mixer in a direct-conversion receiver must be very linear to suppress intermodulation by interferers, and must not let the local oscillator signal leak out of the antenna. Both these problems are circumvented by an unusual, uniquely CMOS solution: a sub-sampling downconversion mixer (Fig.7) [16]. The circuit draws 4 mA to acquire samples of a 900 MHz modulated waveform at a 50 MHz rate which it directly translates to baseband. The linearity, as measured by a third-order intercept of +22 dBm, exceeds that of most continuous-time 900 MHz monolithic mixers.

Key to a frequency-hopped transceiver is a low-power, agile frequency synthesizer with high spectral purity. The requirements of purity and wideband agility are at odds in a VCO-based PLL synthesizer. In this work, a direct-digital frequency synthesizer (DDFS) produces digital words representing accurate samples of a sinewave, whose frequency is set anywhere from DC to half the clock rate by an 11b input control word. A low-power 50 MHz 10b D/A Converter follows to produce a baseband spread-spectrum. A CMOS implementation of the DDFS-DAC synthesizes sinewaves up to 13 MHz with spurious levels lower than –57 dBc while dissipating only 40 mW from 3 V (Fig.8) [17]. A fixed frequency 915 MHz local oscillator, consisting of a four-stage MOS ring oscillator (Fig.9) locked in a PLL to a lower frequency crystal reference, then upconverts this spectrum to the 902-928 MHz band. The PLL bandwidth is principally optimized to suppress oscillator noise. Quadrature outputs at 915 MHz for the image-reject mixers are tapped off at diametrically opposite points to within a phase accuracy of a couple of degrees. The power amplifier consists of a binary-weighted array of FETs biased near threshold, followed by a matching network. FETs in the array are digitally selected to deliver power levels as high as 20 mW to the antenna with a 45% conversion efficiency [18].

Such a "VLSI radio" would represent a major step forward in the gradual evolution of the low-power integrated radio as described in this paper.

1. Hishida, Y. and M. Kazama, *Mobile Communication System: Its Semiconductor Technology*, in *Hitachi Review*. 1993, pp. 95-100.
2. Kasperkovitz, W.G., "FM Receivers for Mono and Stereo on a Single Chip", *Philips Technical Review*, v. 41(6), pp. 169-182, 1983.
3. Abidi, A.A., "Noise in Active Resonators and the Available Dynamic Range", *IEEE Trans. on Circuits and Systems*, v. 39(4), pp. 296-299, 1992.
4. van Dooremolen, W.H.A. and M. Hufschmidt, *A Complete FM Radio on a Chip*, in *Electronic Components and Applications*. 1983, pp. 159-170.
5. Okanobu, T., H. Tomiyama, and H. Arimoto, "Advanced Low-Voltage Single Chip Radio IC", *IEEE Trans. on Consumer Electronics*, v. 38(3), pp. 465-475, 1992.
6. Vance, I.A.W., "Fully Integrated Radio Paging Receiver", *IEE Proceedings*, v. 129, Part F(1), pp. 2-6, 1982.
7. Wilson, J.F., *et al.*, "A Single-Chip VHF and UHF Receiver for Radio Paging", *IEEE J. of Solid State Circuits*, v. 26(12), pp. 1944-1950, 1991.
8. Burt, A., *Direct Conversion Receivers Come of Age in the Paging World*, in *GEC Review*. 1992, pp. 156-160.
9. Tamura, Y., *et al.*, *Development of Advanced Mobile Telephone (Personal Pocket Phone)*, in *NEC Research and Development*. 1990, pp. 60-70.
10. Fenk, J., *et al.*, "An RF Front-End for Digital Mobile Radio", in *Bipolar Circuits and Technology Meeting*, pp. 244-247, 1990.
11. Sevenhans, J., *et al.*, "An Integrated Si Bipolar RF Transceiver for a Zero IF 900 MHz GSM Digital Radio Frontend of a Hand Portable Phone", in *Custom IC Conference*, pp. 7.7/1-4, San Diego, CA, 1991.
12. Leonard, M., *PCMCIA-Sized Radio Links Portable WLAN Terminals*, in *Electronic Design*. 1993, pp. 45-50.
13. Devlin, L.M., *et al.*, "A 2.4 GHz Single Chip Transceiver", in *Microwave & Millimeter-Wave Monolithic Circuits Symp.*, pp. 23-26, Atlanta, GA, 1993.
14. Min, J., *et al.*, "An All-CMOS Architecture for a Low-Power Frequency-Hopped 900 MHz Spread-Spectrum Transceiver", in *Custom IC Conf.*, pp. 16.1.1-16.1.4, San Diego, CA, 1994.
15. Chang, J.Y.-C., A.A. Abidi, and M. Gaitan, "Large Suspended Inductors on Silicon and their use in a 2-μm CMOS RF Amplifier", *IEEE Electron Device Letters*, v. 14(5), pp. 246-248, 1993.
16. Chan, P.Y., *et al.*, "A Highly Linear 1-GHz CMOS Downconversion Mixer", in *European Solid-State Circuits Conf.*, pp. 210-213, Sevilla, Spain, 1993.
17. Chang, G., *et al.*, "A Low-Power CMOS Digitally Synthesized 0-13 MHz Agile Sinewave Generator", in *Int'l Solid State Circuits Conf.*, pp. 32-33, San Francisco, 1994.
18. Rofougaran, M., *et al.*, "A 900 MHz CMOS RF Power Amplifier with Programmable Output", in *Symp. on VLSI Circuits*, Honolulu, 1994.

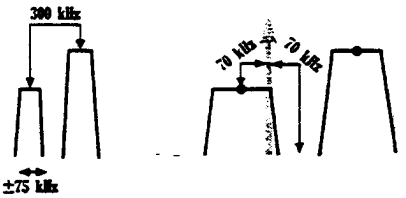


Fig.1. Channel spacing in the broadcast FM band, and the image frequency when the IF is reduced to 70 kHz.

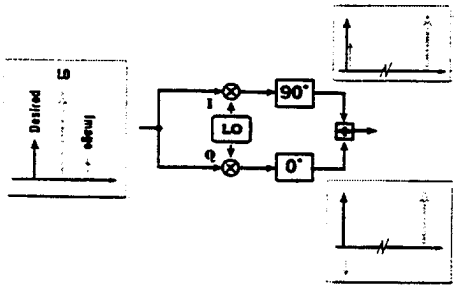


Fig.2. The image-cancelling mixer. This electronic alternative to a highly selective bandpass filter has only become practical on ICs, as the extent of image suppression depends on the matching of the two arms.

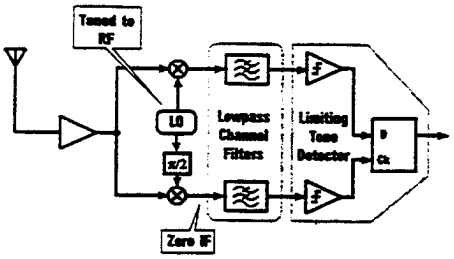


Fig.3. A simple zero IF receiver for binary-FSK signals. Including the lowpass channel filters,this may be readily integrated on a single IC.

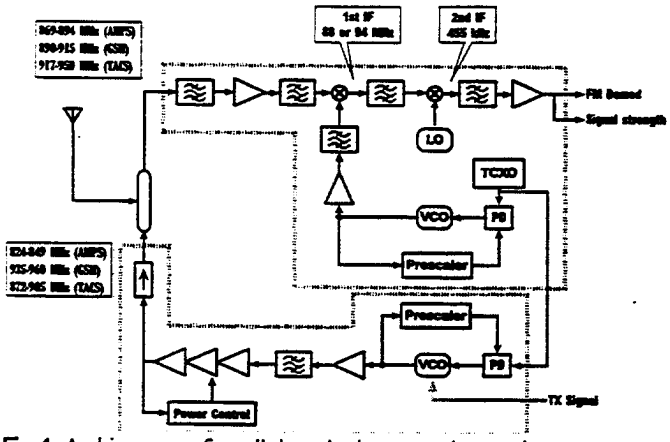


Fig.4. Architecture of a cellular telephone receiver and transmitter. Double conversion affords high selectivity, but at the expense of a large number of discrete filters.

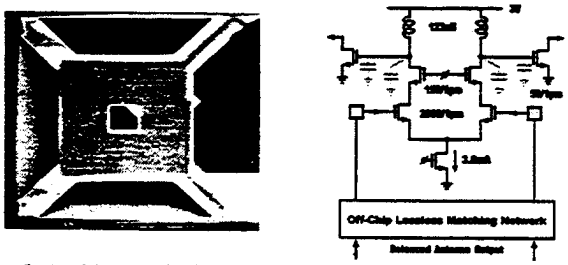


Fig.6. A 100 nH spiral inductor suspended over the silicon substrate. Self-resonance frequency is 3 GHz. A balanced CMOS RF amplifier affords 30 dB gain, 4 dB noise figure at 915 MHz.

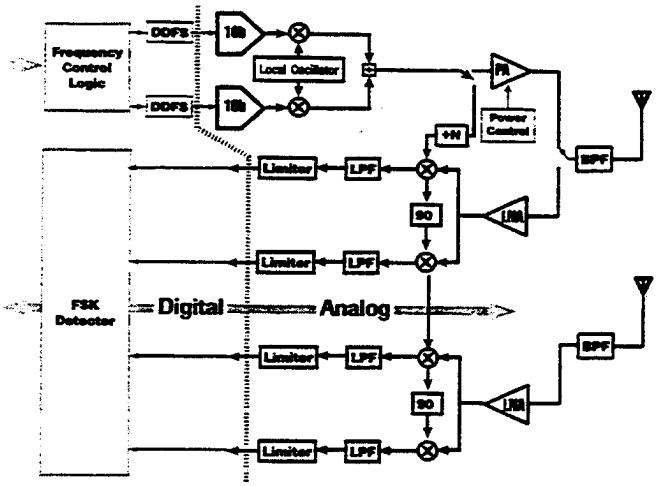


Fig.5. Architecture of the UCLA Frequency-Hopped FSK Transceiver. Two entire receive channels connected to the antennas help to combat multipath fading with spatial diversity.

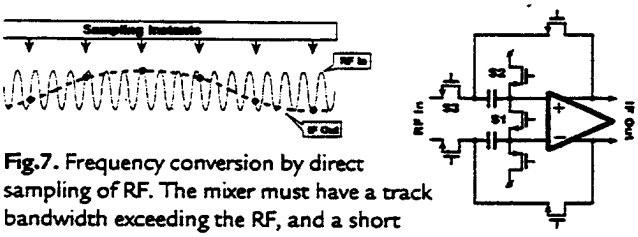


Fig.7. Frequency conversion by direct sampling of RF. The mixer must have a track bandwidth exceeding the RF, and a short sampling aperture.

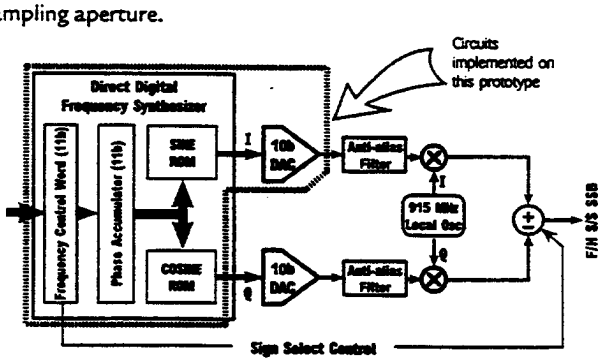


Fig.8. An agile frequency source using direct-digital frequency synthesis. Image reject mixer produces single-sideband, suppressed carrier output. Sign of summation at mixer outputs selects upper or lower sideband, so DDFS need only span 0 to 13 MHz to cover 902-928 MHz.

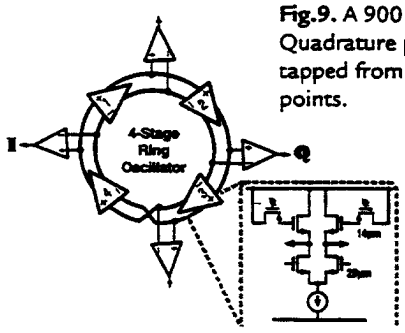


Fig.9. A 900 MHz CMOS VCO. Quadrature phases at any frequency tapped from diametrically opposite points.

Comparison of MOM and FDTD for Radiation and Scattering Involving Dielectric Objects

J.S. Colburn, M.A. Jensen, and Y. Rahmat-Samii

Proceedings of the IEEE Antennas and Propagation Society International Symposium
pp. 1802-1805
June 1994

Comparison of MOM and FDTD for Radiation and Scattering Involving Dielectric Objects †

J. S. Colburn*, M. A. Jensen and Y. Rahmat-Samii
Department of Electrical Engineering
University of California, Los Angeles
Los Angeles, CA 90024-1594

I. INTRODUCTION The influence of dielectric bodies on electromagnetic fields is a key consideration in many practical engineering problems. In many applications, the scattering from dielectric bodies for known incident fields is of interest. In other situations where a dielectric object interacts with a radiator - such as biological tissue near a wireless transceiver or a dielectric load on an antenna - the resulting input impedance, radiation patterns and power absorption characteristics are important quantities to determine. Simulation of such configurations often requires the use of computational tools such as the method of moments (MOM) or finite difference time domain (FDTD) approach. Because these techniques are based on different solution methodologies, they provide an independent check of computed results. This paper will provide a comparative study of these two approaches. Results from various scattering and radiation examples will be presented, and the advantages of each technique discussed.

II. FORMULATION The method of moments is a commonly used technique to convert linear integral formulations into a system of equations which can be solved numerically. A volume integral equation can be written to account for the presence of a dielectric object. The dielectric body is replaced with an equivalent free-space current density, which is then used to express the scattered field inside the body using the free-space Dyadic Green's function. For a known incident field, the resulting integral equation can be solved numerically using the method of moments to determine the equivalent free-space current density inside the body [1]. With this current density determined, the scattering from the dielectric body can be calculated using a far-field integration. For problems involving radiation from thin-wire antennas, Pocklington's thin-wire formulation can be used. Coupling between the dielectric object and the antenna is accounted for by forcing the incident field in the body to be that radiated by the antenna and applying proper boundary conditions. This results in a set of coupled integral equations to be solved simultaneously.

The finite difference time domain approach is a numerical technique used to solve Maxwell's equations in the time domain. This methodology uses a spatial and temporal discretization of Maxwell's differential equations to determine the time-evolution of the fields within a body subjected to electromagnetic radiation [3].

III. RESULTS Computer programs were written to implement both the MOM and FDTD approaches to solve this class of problems. In the MOM implementa-

†This work was supported by ARPA Contract #DAAB07-93-C-C501.

tion, a $(3N_b + N_a)$ system of equations results, where N_a is the number of basis functions used for the wire antenna and N_b is the number of subvolumes the dielectric body is divided into. Attention was given to the convergence of the matrix elements. The relative location and orientation of the dielectric object and wire antenna can be arbitrary, with no effect on the memory required to solve the problem. The FDTD approach was implemented on a rectangular grid, which must encompass both the radiator and the dielectric body. Thus with this technique, the relative location and orientation of the radiator and dielectric body is restricted. The memory required for this technique grows proportional to $6N_n$, with N_n being the number of nodes in the grid, whereas for the MOM technique the memory required grows proportional to $(3N_b + N_a)^2$. Thus for a given sized memory, the FDTD approach can handle more elaborate models in most cases.

The first set of results shown, Figure 2, is the normalized far-field scattered from a homogeneous 0.2 λ_o dielectric cube with a relative dielectric constant of 4.0. The incident field is a x-polarized, positive z-propagating plane wave. From Figure 2, one can see the results from these two approaches compare very well for this simple case. Also plotted on Figure 2 is computed data from [4].

Figure 3 illustrates a 3 dimensional structure made up of seven, 0.1 λ_o lossy homogeneous cubes, each with a relative dielectric constant of $10 - j10$. Figure 4 gives the calculated scattered far-field from the dielectric body seen in Figure 3. The results can be seen to compare well, with the exception of the null depth in the computed scattered far-field.

Next consider the perturbation (Z_p) of the input impedance (Z_{in}) for a dipole near a dielectric body from that in free space, $Z_p = Z_{in} - Z_o$, where Z_o is the free space input impedance of the dipole. Figure 5 is a schematic of an example case run, where the dielectric object has the material properties 0.5 N saline solution at 600 MHz, $\epsilon_r = 71.0$, $\sigma = 4.4$ S/m and density $\rho = 1000$ kg/m^3. Figure 6 is a plot of Z_p as a function of the distance between the dipole and the center of the dielectric body (d). Also plotted in Figure 6 is measured data from [5]. In the MOM implementation a magnetic frill source model was used and in the FDTD implementation the delta gap excitation was applied. From these results good correlation can be seen between the two approaches.

Figures 7 (a) and (b) are plots of the computed specific absorption rate (SAR = $\frac{\sigma |E|^2}{2 \rho}$) for the x-y and x-z planes respectively with d = 15 cm. Figure 8 shows the computed SAR (dB) for the y-z plane, also for d = 15 cm. Differences in the results obtained by the two methodologies are attributed to the differences in the geometrical gridding used in each technique as specified in the figure captions.

REFERENCES

[1] D. E. Livesay and K. Chen, *IEEE Trans. on Microwave Theory and Tech.*, 22:1273-1280, 1974.

[2] K. Karimullah, K. Chen and D. P. Nyquist, *IEEE Trans. on Microwave Theory and Tech.*, 28:1218-1225, 1980.

3. K. S. Yee, *IEEE Trans. Antennas Propag.*, 14:302-307, 1966.
4. T. K. Sarkar, E. Arvas and S. Ponnapalli, *IEEE Trans. Antennas Propag.*, 37:673-676, 1989.
5. K. Karimullah, PhD dissertation 1979, Michigan State Univ.

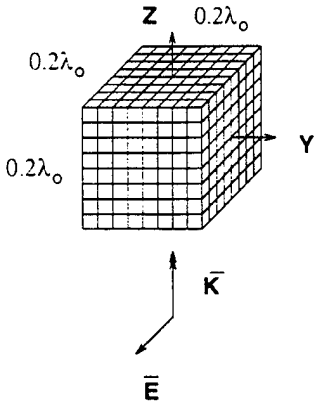


Figure 1: 0.2 λ_o cube, ϵ_r = 4.0

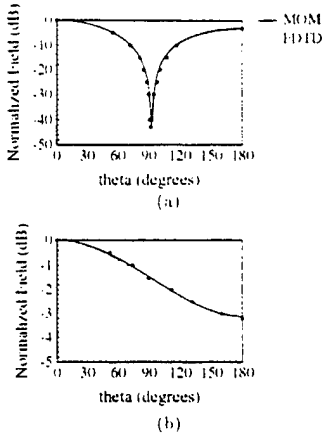


Figure 2: Normalized Scattered Field (a) $|E_\theta|$, $\phi = 0°$ cut (b) $|E_\phi|$, $\phi = 90°$ cut; dots: computed [4]

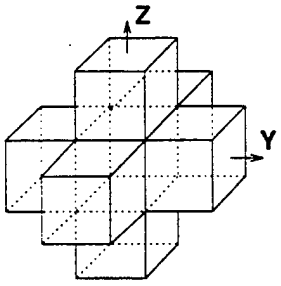


Figure 3: Seven 0.1 λ_o cubes, ϵ_r = 10 - j 10

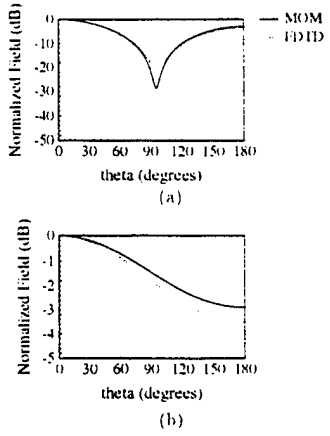


Figure 4: Normalized Scattered Field (a) $|E_\theta|$, $\phi = 0°$ cut (b) $|E_\phi|$, $\phi = 90°$ cut

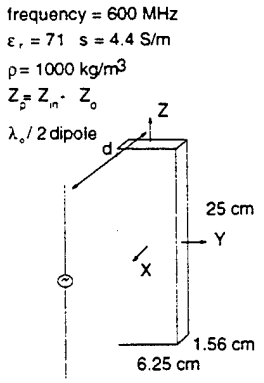


Figure 5: Schematic of Impedance Perturbation Example

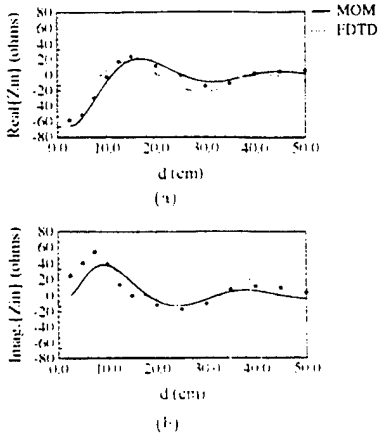


Figure 6: Perturbation of (a) Real$\{Z_p\}$ (b) Imag $\{Z_p\}$: dots, measured [5]

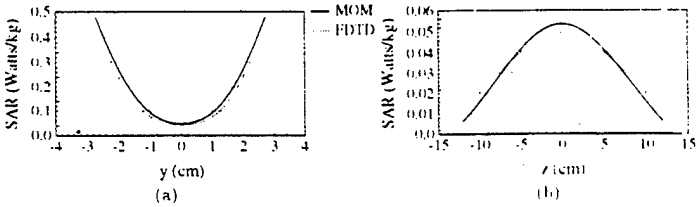


Figure 7: SAR for d = 15 cm (a) x-y plane: MOM z = 0.39 cm, x = 0.39 cm, FDTD z = 0 cm, x = 0.5 cm (b) x-z plane: MOM y = 0.39 cm, x = 0.39 cm, FDTD y = 0.5 cm, x = 0.5 cm

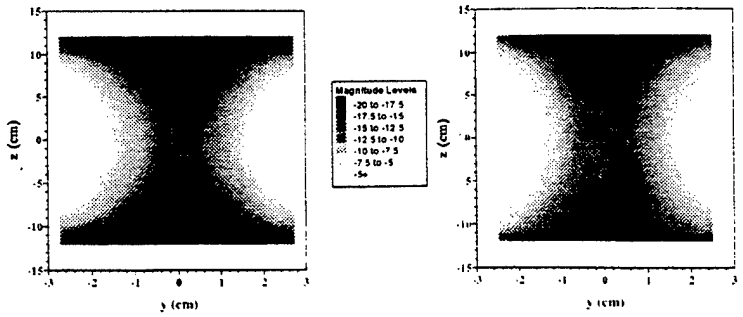


Figure 8: SAR (dB) for d = 15 cm in the y-z plane (a) MOM x = 0.39 cm (b) FDTD x = 0.5 cm

The Electromagnetic Interaction Between Biological Tissue and Antennas on a Transceiver

M.A. Jensen and Y. Rahmat-Samii

Proceedings of the IEEE Antennas and Propagation Society International Symposium
pp. 367-370
June 1994

The Electromagnetic Interaction Between Biological Tissue and Antennas on a Transceiver Handset †

Michael A. Jensen* and Yahya Rahmat-Samii
Department of Electrical Engineering
University of California, Los Angeles
Los Angeles, CA 90024-1594

I. INTRODUCTION In the design of antennas for use with hand-held transceivers, the electromagnetic interaction between the antenna and the nearby biological tissue is a key factor which must be considered when selecting the antenna configuration. The engineer must consider the effect of the body on the antenna performance as well as the rate at which energy is absorbed in the tissue. Investigation of such issues can be performed through numerical simulation provided that the antenna, handset, and adjacent tissue are represented with detailed models and that accurate solution methodologies such as the finite-difference time-domain (FDTD) approach are used. Past studies have appeared which use the FDTD method either with simple models of the head (e.g. a homogeneous sphere) [1] or simple antenna models [2] to predict the interaction between the antenna and the nearby tissue. This paper combines detailed models of the head and hand with real-life models of antennas on transceiver units to predict the input impedance, radiation patterns, gain, and Specific Absorption Rate (SAR) which result when the handset operates in the proximity of a human. Results are shown for both the monopole and planar inverted F antenna (PIFA) to illustrate the interaction for external and internal antennas respectively.

II. MODELING The two different antenna/handset configurations used in this investigation are depicted in Figure 1. Figure 1(a) illustrates the geometry for a monopole antenna on a conducting chassis. Figure 1(b) represents the topology for the side-mounted PIFA's arranged in a diversity fashion. The handset chassis dimensions shown in (a) are valid for both configurations.

Biological tissue is modeled using computational phantoms with inhomogeneous electrical properties which represent the different tissues in the hand and head. The hand consists of a layer of bone surrounded by a layer of muscle. The head phantom is composed of the seven tissues listed in Table 1. The location of the various tissues has been determined from magnetic resonance images as well as cross-sectional anatomy books. The phantoms are shown with the handset in Figure 1(c). The electrical parameters for each type of tissue are provided in Table 1 [2].

III. RESULTS For both antennas, the wires have a diameter of 0.9195 mm. A wire-subcell method is used to account for this finite wire radius in order to obtain accurate values for the input impedance. The value of $|S_{11}|$ (assuming a 50Ω feeding coax) versus frequency is shown in Figure 2(a) for the monopole on

†This work was supported by ARPA Contract #DAAB07-93-C-C501.

0-7803-2009-3/94/$4.00 © 1994 IEEE.

the handset. Computed and measured data is provided for scenarios where no tissue is present, where the hand only is present, and where both the head and hand are present. As can be seen, the presence of the head and the hand exercises relatively little influence on the input impedance of this antenna. Also noteworthy is the close correspondence between the measured and computed results.

As might be expected, the placement of the hand exercises considerably more influence on the impedance of the PIFA. This is illustrated in Figure 2(b). In this configuration, to simulate the fact that a plastic case separates the hand from the antenna, a 6.56 mm layer of air is maintained between the hand and the handset. The different curves in Figure 2(b) represent the results for no hand and for the hand at three positions, with the distances representing the distance from the top of the hand to the top of the handset. As the hand begins to cover the antenna, significant detuning of the antenna occurs. This is a key issue to consider before placing integrated antennas on the handset. The measured data, provided for two of the cases, again shows very good correlation with the computed results.

The radiation patterns normalized to the antenna gain for both the monopole and the side-mounted PIFA are provided in Figure 3. In both cases, the handset is rotated 60° from upright in the $-y$ direction to simulate the position commonly assumed while speaking. In Figure 3(a), the patterns for the monopole on the handset with no operator present are compared to the patterns when the hand and head are present. As can be seen, the interaction with the tissue results in noticeable changes to the shape, polarization, and gain characteristics of the pattern. Figure 3(b) illustrates the patterns with the operator present for the PIFA antenna. These patterns also show the reduced gain resulting from absorption in the tissue.

Table 2 provides a quantitative presentation of the amount of power absorbed in the hand and head as well as the antenna efficiency η_a. In these ratios, P_{abs} is computed from the integral

$$P_{abs} = \frac{1}{2}\int_V \sigma |\mathbf{E}|^2 dV$$

over the volume of the tissue where σ is the tissue conductivity. P_{tot} represents the total power delivered to the antenna. This table also provides the maximum values of the $SAR = \sigma|\mathbf{E}|^2/2\rho$ in the head and hand, where ρ is the tissue density. These values are averaged over 1 gram of tissue (assuming 0.6 W of delivered power). The IEEE limit for the peak SAR to any 1 gram of tissue for 30 minutes or more is 1.6 W/kg. Comparisons such as those presented in Table 2 are very useful for determining the suitability of different radiators for personal communications applications.

REFERENCES

[1] J. Toftgard, S. N. Hornsleth, and J. B. Andersen, *IEEE Trans. Antennas Propag.*, 41:739–746, 1993.

[2] P. J. Dimbylow and O. P. Gandhi, *Phys. Med. Biol.*, 36:1075–1089, 1991.

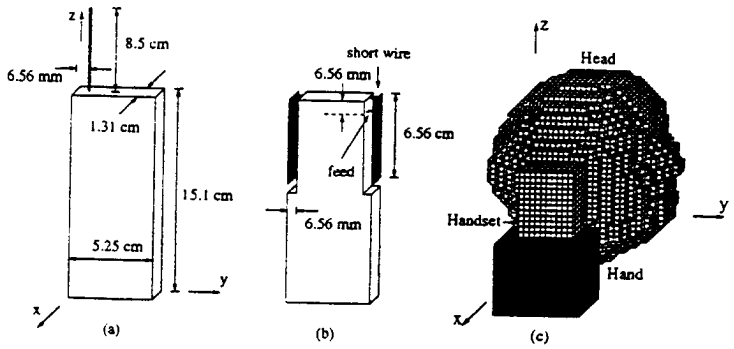


Figure 1: Geometries and dimensions for (a) the monopole on the handset, (b) the side-mounted PIFA on the handset, and (c) the hand and head.

Table 1: Relative permittivity, conductivity, and density of the tissues in the hand and head near 900 MHz.

Tissue	Permittivity	Conductivity (S/m)	Density ($\times 10^3$ kg/m^3)
Bone	8.0	0.105	1.85
Skin	34.5	0.60	1.10
Muscle	58.5	1.21	1.04
Brain	55.0	1.23	1.03
Humour	73.0	1.97	1.01
Lens	44.5	0.80	1.05
Cornea	52.0	1.85	1.02

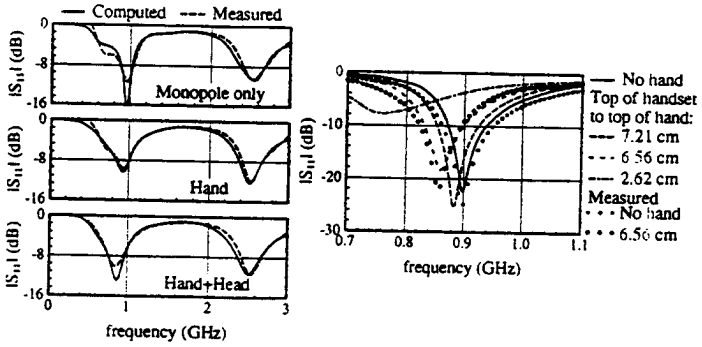
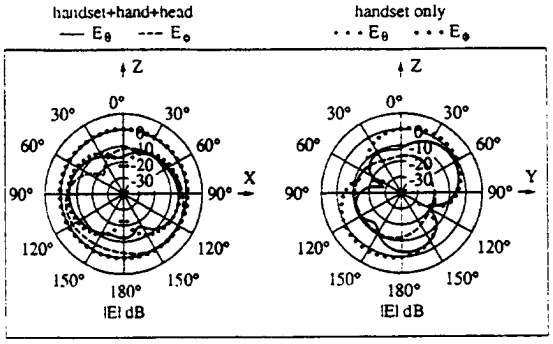
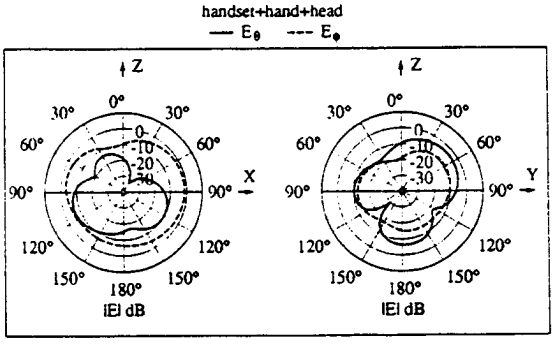


Figure 2: $|S_{11}|$ versus frequency for the antennas on the handset: (a) monopole with no tissue, with hand, and with hand and head; (b) PIFA with no tissue, with the hand at three positions. Measured data is shown for several cases.



(a)



(b)

Figure 3: Gain patterns at 915 MHz for the (a) monopole and (b) PIFA on the handset with the hand and head present. The handset is rotated 60° from upright toward the $-y$ direction.

Table 2: Percentage of the power absorbed in the head and head, the radiation efficiency, and the peak SAR (for 0.6 W power delivered to antenna) in the head and head for the monopole and PIFA configurations. All data is computed at 915 MHz.

Configuration	P_{abs}^{head}/P_{tot}	P_{abs}^{hand}/P_{tot}	η_a	SAR_{max}^{head} (W/kg)	SAR_{max}^{hand} (W/kg)
Monopole	0.340	0.188	0.472	1.23	1.45
Side Mounted PIFA	0.335	0.325	0.340	1.75	4.10

A 900 MHz CMOS RF Power Amplifier with Programmable Output

M. Rofougaran, A. Rofougaran, C. Olgaard, and A.A. Abidi

1994 Symposium on VLSI Circuits Diges of Technical Papers
pp. 133-134
June 1994

A 900 MHz CMOS RF Power Amplifier with Programmable Output

M. Rofougaran, A. Rofougaran, C. Ølgaard, and A. A. Abidi

Integrated Circuits & Systems Laboratory
Electrical Engineering Department
University of California
Los Angeles, CA 90024-1594

Introduction

The power amplifier module constitutes the largest current drain on a wireless transceiver during transmit mode. In future cellular networks for digital wireless transceivers, the base station will adaptively regulate the transmitted *power levels* of each transceiver to enable the largest possible number of users to share a wireless channel. This requires a high-efficiency power amplifier with a digitally selectable output level spanning a wide range. The power amplifier reported here is intended for use in an all-CMOS frequency-hopped spread-spectrum transceiver operating in the 902-928 MHz band. It delivers a controllable power between 20-μW and 20-mW from a 3-V supply to the antenna.

Circuit Design

To deliver 20-mW into a typical 50 Ω antenna load, the power amplifier must apply a single-ended voltage swing of 2.8 V peak-to-peak across the load. With a 3-V power supply, this requires the use either of an impedance transformer, or of a differential drive to the antenna. The latter is used here. The quasi-differential power amplifier consists of a preamplifier stage, followed by a binary-weighted array of driver FETs selected by switches. The balanced output of an upconversion mixer drives the two identical halves of the power amplifier.

For high power-conversion efficiency, the driver FETs are biased in Class-C mode close to cutoff at a V_{GS}-$V_t \approx 200$ mV. This bias point was chosen after extensive simulations, to balance the specifications on maximum delivered power and the capacitance of the driver FETs. The gate-bias of these FETs is adjusted by an adaptive common-mode bias control voltage, which determines the level shift in a source follower between the preamplifier and driver FETs (Fig.1). An inductor load biases the preamplifier output to the supply voltage of 3-V, and allows the signal swing to exceed the supply. This on-chip inductor serves three vital functions. First, by tuning the capacitances of the 1-μm MOSFETs, it peaks the gain about 900 MHz. Second, it makes it possible to deliver a large signal to the driver stage. Third, as the signal swing exceeds the supply, it compensates for the source follower V_{GS} voltage drop which is exacerbated by the body effect on NMOSFETs. The source follower width is scaled with the driver FET width. A PMOSFET switch in series with the drain enables each source follower, and the corresponding driver FET. This method of switching least loads the high-frequency signal path. A five-element array of binary weighted source followers and driver FETs, whose inputs and outputs in parallel, are connected to one preamplifier (Fig.2). The preceding upconversion mixer, to be integrated on-chip, will drive the preamplifier with a constant envelope sinewave. A 5-bit word applied to the PMOSFET select switches sets the power delivered to the antenna load across a 30 dB range.

The off-chip lossless matching network (Fig.2) between this IC and the antenna was an integral part of the power amplifier design. A high-frequency chip inductor tunes the output capacitance of the driver FET array and the pad capacitance of the IC for maximum power delivery into a 50 Ω load. This will also form a lowpass filter above the 902-928 MHz band of operation to attenuate harmonics created by the Class-C power amplifier. A three pole-pair dielectric resonator bandpass filter with a 26 MHz passband centered at 915 MHz is inserted between the amplifier output and antenna to suppress out-of-band noise and spurious components, while precisely defining the band of transmission. In a frequency-hopped system, only one tone is applied to the power amplifier at a time, so intermodulation products are negligibly small.

Experimental Results and Discussion

The power amplifier was fabricated in a 1-μm CMOS process, and including two on-chip spiral inductors it occupies about 1 sq mm of active area (Fig.3). The IC was mounted in a microwave package, in very close proximity to the two chip inductors used in the output matching network. For purposes of standalone testing, the inputs of the circuit were terminated on-chip with 50 Ω resistors, although in the complete system a voltage will drive the input. The two differential outputs were combined in a balun for single-ended measurements. In effect, this corresponds to the power amplifier driving a 100 Ω balanced antenna, such as a loop.

Measurements of the frequency response show that the on-chip inductor, which does not embody the sharp tuning characteristics of the suspended inductors we have reported previously [1], provides a wide gain characteristic nominally centered around 900 MHz. The off-chip matching network followed by the dielectric resonator filter precisely defines the transmit channel (Fig.4). With a constant input voltage, the output power may be swept over the expected 30 dB range under control of the digital word (Fig.5); power-level switching is essentially instantaneous. The power conversion efficiency of the amplifier approaches the designed 30% at maximum power, but drops off at low power levels where it is in any case unimportant. This efficiency is comparable to what has been reported

Research supported by ARPA, Rockwell International and the State of California MICRO Program

0-7803-1918-4/94/$3.00 © 1994 IEEE

in other power amplifiers fabricated in more complex technologies operating at 900 MHz or 2 GHz [2-4]. At 1-dB compression, this prototype delivers 32 mW (15 dBm) (Fig.6).

The first 900 MHz CMOS power amplifier operating at 3 V is reported. It exploits unique features of standard CMOS to provide digitally selectable power control over three orders of magnitude, to attain Class-C operation, and it exploits on-chip inductors to obtain high-frequency performance. This is intended to be integrated with an upconversion mixer and its preceding components as part of an all-CMOS wireless transceiver.

1. Chang, J.Y.-C., A.A. Abidi, and M. Gaitan, "Large Suspended Inductors on Silicon and their use in a 2-micron CMOS RF Amplifier", *IEEE Electron Device Letters*, v. 14, pp. 246-248, May 1993.
2. Yoshida, L, *et al.*, "Highly Efficient 1.5 GHz Si Power MOSFET for Digital Cellular Front End", in *Int'l Symp. on Power Semiconductor Devices*, pp. 156-157, Tokyo, 1992.
3. Kermarrec, C., *et al.*, "High Performance, Low Cost GaAs MMICs for Personal Phone Applications", in *Int'l Symp. on GaAs and Related Compounds*, Karuizawa, Japan, 1992.
4. Murai, S., *et al.*, "A High Power-Added Efficiency GaAs Power MESFET Operating at a Very Low Drain Bias for use in L-Band Medium-Power Amplifiers", in *GaAs IC Symposium*, pp. 139-142, 1992.

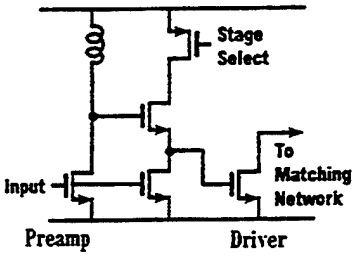


Fig.1. Single stage (half of quasi-differential) of power amplifier, showing preamplifier, level-shift, and output driver. Level shift is controlled by input common-mode.

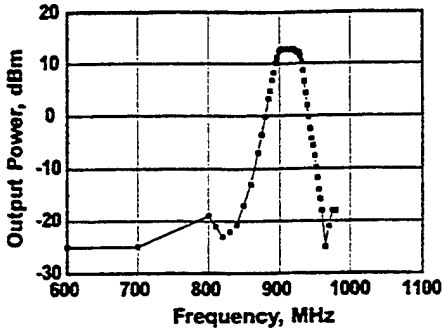


Fig.4. Measured frequency response of output power. Matching network and off-chip dielectric resonator filter define channel selectivity.

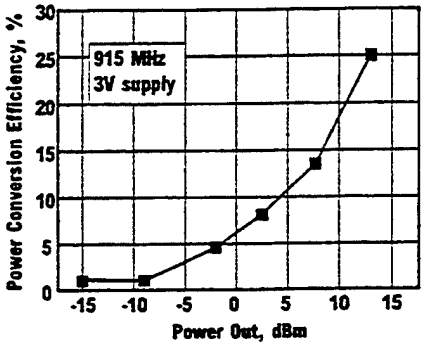


Fig.5. Measured power output vs. digital selection, at constant input voltage, shows sought 30 dB range and efficiency approaching 30%. Efficiency is the ratio of power delivered to load to the power drawn from the supply.

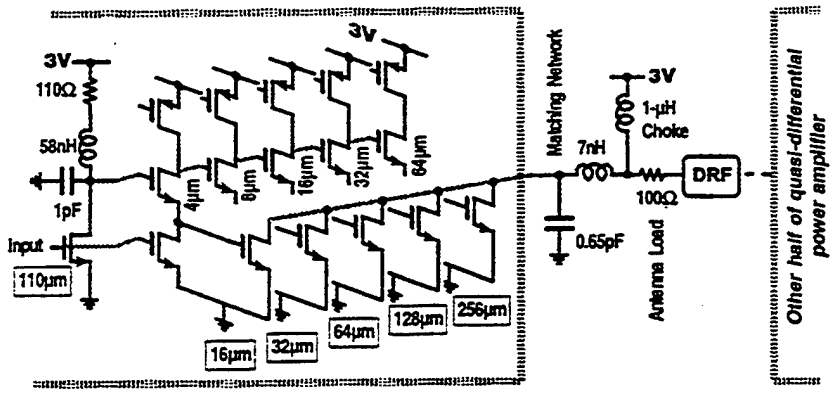


Fig.2. One half of complete power amplifier. Inputs and outputs of binary-weighted array of FETs are in parallel; FETs are selected by PMOSFET switches. Antenna is driven differentially by two such circuits. DRF is dielectric resonator filter. Bond pad capacitance and wirebond inductance are part of the matching network.

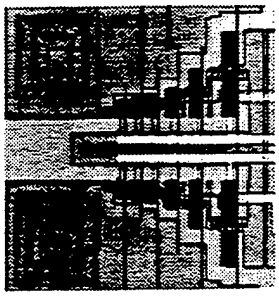


Fig.3. Detail of chip microphotograph. On-chip 60 nH inductors are 0.25 mm on a side.Outputs of binary-weighted array of driver FETs in quasi-differential amplifier are merged.

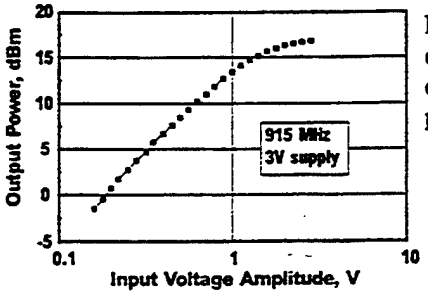


Fig.6. Measured compression characteristic of power output.

Radio-Frequency Integrated Circuits for Portable Communications

A.A. Abidi

Proceedings of the IEEE 1994 Custom Integrated Circuits Conference
pp. 151-158
May 1994

Radio-Frequency Integrated Circuits for Portable Communications

Asad A. Abidi
Electrical Engineering Department
University of California, Los Angeles

Introduction

While the greatest design efforts in silicon analog ICs have been historically aimed at baseband circuits, which range in frequency of operation from voiceband to communications circuits at tens of MHz, analog GaAs circuits have found a niche in microwave communications ICs by exploiting the high-frequency capabilities of the MESFETs and the semi-insulating properties of the substrate material to form on-chip inductors and transmission lines. These monolithic microwave ICs (MMICs) have enabled entirely new applications in the microwave and millimeter-wave ranges, from 10 GHz to beyond 100 GHz.

A new market has meanwhile emerged in the form of portable wireless communications devices, operating in the 900 MHz to 2 GHz range, where miniaturization and low-energy operation is sought through the aggressive large-scale integration typical of silicon ICs, yet the circuit design style at times resembles MMICs. Early indications suggest that silicon will be the technology of choice in this application, because analog and digital functions can freely intermingle on the IC, and because the transistor f_T both in production bipolar and CMOS technologies is adequate for operation in this range. Although there are only a few examples today of these "VLSI radios", they arouse much curiosity and speculation because it is anticipated that by the year 2000, close to 300 million of these consumer wireless devices will be in use [1].

This paper summarizes some of the key radio-frequency integrated circuit developments in the last thirty years, the majority of which have been developed for consumer use. From this survey, it projects what developments may be reasonably expected in the architectures and circuit blocks of RF ICs over the next five to ten years.

The UCLA Wireless Transceiver Project is sponsored by the Advanced Research Projects Agency, and by a consortium of semiconductor companies under the State of California MICRO Program.

ICs in Broadcast Radio Receivers

Consumer electronics companies have been drawn, since the early days of IC technology, to the prospect of miniaturizing broadcast AM/FM receivers for use in the home, car, or even in the pocket. Following early efforts in the 1970's, AM/FM radios were reduced to either a single-chip containing functions from the IF amplifier onwards to the audio power amplifier [2], or to two- or three-chip sets [3, 4], surrounded by a dozen or more discrete filter modules, consisting of *LC* filters and, at high frequencies such as 10.7 MHz, ceramic resonator filters. ICs enabled unique solutions to some baseband functions, such as cheap phase-locked loops for the detection of pilot tones or for FM demodulation. In the front-end, the superior characteristics of the double balanced mixer using the four-quadrant analog multiplier could only be realized in IC form, while the RF tuned amplifiers and most of the inter-stage coupling elements remained discrete. However, ICs did not influence the architecture and frequency allocation plan in these receivers, which continued to resemble conventional discrete-component radios.

A re-thinking of the conventional FM receiver came about at Philips where Kasperkovitz [5, 6] realized that the main advantages of integration would come about by integrating the many filters on-chip, thus eliminating the multiple trims of the discrete filters in production. However, it was then impossible to integrate the RF or the 10.7 MHz IF bandpass filters on-chip either with spiral monolithic inductors or by simulating them electronically with gyrators. The key step in solving this problem was to lower the intermediate frequency from 10.7 MHz to 70 kHz. A small off-chip inductor provided some RF filtering, but the channel bandwidth was mainly defined on-chip by high-order active *RC* filters. The power dissipation in these filters for a given dynamic range scales down with the IF. This choice of IF also meant that the image frequency lies half way to the next FM channel (Fig. 1), so no special effort was required to suppress the image. As a consequence of the low IF, the received frequency deviation of ± 75 kHz was compressed to ± 15 kHz by a negative frequency feedback on a varactor-tuned VCO to avoid distortion caused by signal aliasing around DC. Integration also

0-7803-1886-2/94 $3.00

made it cost-effective to use a new muting system. This single-chip FM mono receiver [7] requires only 15 small capacitors and two inductors, making it possible to fit the entire system within a wristwatch. The 8 mA drain from a 4.5V supply made the bipolar IC attractive for use in miniature portable applications requiring a small battery volume.

The same idea of low IF has been pursued at Sony, seeking bipolar ICs operating at supplies as low as 0.95 V provided by a single battery [8]. It is likely that this research program has contributed to their now universally recognized product line of personal entertainment systems. Their latest FM radio IC is a double superheterodyne architecture. At 30 MHz, the first IF is high enough that a fixed bandpass filter in series with the antenna spanning the entire broadcast range from 80 to 110 MHz will substantially suppress the image signal; no on-chip filter is used. Then, a second mixer converts to a low IF of 150 kHz, and thereafter on-chip filtering is implemented by a 9th order active *RC* lowpass section. This IF is high enough to include the adjacent channel in the image, so an image rejection mixer (Fig.2) is used. This electronic solution to a difficult filtering problem was first described in the 1930's, but has only become practical on the IC. The image suppression is limited by matching in the two channels to 40 dB. An allpass active *RC-CR* filter produces quadrature phases of the local oscillator signal. This receiver drains about 15 mA in either FM or AM mode. The RF amplifier transistors are integrated on these ICs, although they require off-chip inductors for the load, and the local oscillator needs an off-chip *LC* tuned circuit.

ICs in Wireless Paging Receivers

The desirability of paging people on the move through a miniature wireless receiver was recognized in the early 1960s. The Bell System's Bellboy experiments [9] carried out in pre-IC days anticipated many of the key concepts that in later years were used in integrated paging receivers. The prototype system operated at 150 MHz, and frequency-modulated three tones on the carrier which the intended receiver could recognize by the simultaneous response of passive reed resonators. The receiver operated at 4 V using a total of only ten transistors [10]. The low IF of 6 kHz meant that simple capacitively-coupled 10 kHz *lowpass* (rather than bandpass) channel filters could be used, and that transistors after the RF section required small bias currents. A cascode two-stage amplifier in the RF front-end re-uses the same bias current, a concept found even today in low-power MMICs.

Paging receivers have been in continuous development since, evolving into today's wide-area pagers with digital signalling. The modern pager signals anuwehere from 500 to 1200 b/s by encoding data with a simple positive or negative offset of the carrier frequency — the binary frequency-shift keyed (FSK) modulation. Surprisingly, paging receivers up to the early 1980's did not exploit the powerful simplifications this signaling scheme implied, even though there was great market pressure to reduce the battery drain and miniaturize the receiver. A conventional double superheterodyne receiver of that era consisted of a first upconversion of the received signal to suppress the image channel with a crystal bandpass filter of modest complexity, and then a downconversion with further selection with a ceramic bandpass filter, followed by a frequency discriminator to demodulate the FSK [11]. Vance at Standard Telecommunication Laboratories realized how to simplify an FM receiver on to one chip with a quadrature downconversion mixer combined with zero IF. This was exactly the right solution for the binary FSK digital paging receiver [12]. This remarkably simple receiver (Fig.3) consists only of a quadrature demodulator, lowpass filters in each arm, limiters, and a D-type flip-flop detector, and when integrated on one bipolar chip, it drains 2.5 mA from 1.8 V when active, although in standby the drain falls to a mere 50 μA. Large value off-chip capacitors were required for AC coupling and for the lowpass filters. A second low-frequency digital CMOS IC performed all the user interface functions. Zero IF meant that the mixer has *no* image, and because data is encoded in frequencies ±4.5 kHz away from the carrier, DC offsets in the receiver or flicker noise below this frequency could be removed by capacitive coupling in the receiver. The 10 kHz lowpass filters serve to limit out of band noise and blocking signals from degrading the signal-to-noise ratio at the detector.

Pager ICs at Philips initially used a frequency-offset receiver principle [13], whereby the local oscillator frequency is offset from the received carrier by 2 kHz, thus converting the FSK tones to 2.5 kHz and (aliased to) 6.5 kHz. This requires a fairly sophisticated automatic frequency control, and a frequency discriminator to detect the data. Aside from the local oscillator crystal, three off-chip tuned circuits must also be provided. The tuned circuits are obviated in a zero-IF receiver [14], which includes the filters on-chip. These consist of a third-order active *RC* lowpass filter, followed by a 7th order gyrator-based lowpass filter with a 15 kHz cutoff. The signal is amplified at these low frequencies. It is interesting to note that the only components on the front-end not integrated are the tuned load for the RF amplifier transistors,

and the quadrature phase-shift network for the image reject downconversion; on this IC, both are combined in one off-chip signal path.

Others, too, have developed similar zero-IF bipolar integrated front-ends, sharing the feature that much of the die area is taken up by the capacitors for the on-chip lowpass filters [15, 16]. Zero IF and pager signalling were a perfect match, and the RF sections of the wireless paging receiver seems to have evolved little since Vance's work in 1982.

ICs in Cellular Telephone Transceivers

Mobile and handheld cellular telephones are the first widespread two-way radios for consumer use, for which cordless telephones were only a prelude. Radio telephones to be successful with consumers must meet stringent demands for low weight and volume, long battery life, and low cost. Further, to support a large numbers of users in a crowded radio spectrum, radio telephones require more signal processing than other common transceivers. They must therefore use higher levels of integration in the electronics.

The first ICs for use in portable communication devices contained the frequently occurring building blocks of conventional single or double superheterodyne receivers, such as the mixer and local oscillator, or the IF amplifier chain and signal-strength indicator [17]. These silicon ICs would typically be preceded by a discrete RF amplifier and a first mixer. The front-end components themselves, such as a tuned RF low-noise amplifier in the 900 MHz band, a mixer, and a local oscillator, were first integrated on GaAs ICs [18-21]. Aside from the superior high-frequency characteristics of the MESFETs, these ICs included spiral inductors as loads and as series feedback elements for low-noise input matching (Fig.4) [22]. It was by then well known after many years of GaAs MMIC development that the semi-insulating substrate and gold metallization were well-suited to the monolithic fabrication of large value inductors with a modest-Q and a high frequency of self-resonance. Other chip sets have addressed the frequency synthesizer and power amplifier control module [23], or the downconversion and upconversion mixers for the receiver and transmitter, respectively, integrated with their shared local oscillator on the same GaAs substrate [24]. The four-FET switch mixer (Fig.5) used on many GaAs ICs may be very linear, but it requires large local oscillator levels to turn the switches on and off, and unlike the viable alternative of the analog multiplier, it is lossy and will degrade system noise figure.

Conventional cellular telephones typically use the double conversion architecture (Fig.8), assembled with these building blocks, often using GaAs ICs at the receiver and transmitter modules, silicon building blocks in the IF, and surrounded by an impressive collection of discrete resonator filters and antenna duplexers [25, 26]. Very small handsets [27] weighing less than 230g, and occupying a volume less than 150cc, owe more to the use of miniature discrete filters [28] and tiny IC packages [29] than to higher levels of integration in the electronics.

The new generation of digital mobile telephones, with their complex modulation formats and the greater requirements to withstand nearby blocking signals, are forcing yet greater integration of the RF and IF electronics. In a 900 MHz transmitter and receiver silicon bipolar chipset from Siemens [30], a precision quadrature upconversion mixer implements the phase-shift keyed modulation specified in the European GSM system to produce a single-sideband, suppressed carrier output. Including an output stage to drive a separate power amplifier, the IC drains 40 mA. The receiver IC includes a first mixer, with buffers to drive an IF off-chip bandpass SAW filter, and a quadrature downconversion mixer to detect the single-sideband input at baseband. Including an AGC with 70 dB range operating at the 40 to 90 MHz IF and a signal-strength indicator, the receiver IC drains 24 mA. Although the on-chip gain may be as large as 80 dB, it is distributed in different frequency bands, leading to stable operation with little on-chip crosstalk .

In an effort towards even greater miniaturization at Alcatel, the transmitter and receiver sections co-exist on one silicon bipolar chip, in this case with a zero IF architecture [31]. A separate low noise amplifier block drives directly into quadrature (I-Q) mixers with a wide enough dynamic range to withstand large blocking signals only 3 MHz away from the carrier without producing significant intermodulation. The receiver also requires large (650 pF) off-chip capacitors in an anti-alias filter, after which the quadrature downconverted signals are sent to a baseband CMOS IC containing a high-order switched capacitor lowpass filter to remove the blocking signals, and a demodulator. An *RC* and *CR* network with off-chip trimming shifts the local oscillator phase by $\pm 45°$ to produce the quadrature drive to the mixers. There are separate transmission upconversion mixers to drive a power amplifier module. The chip drains 25 mA from 5V in the receive mode, and 45mA in the transmit mode, in this respect rather similar to the Siemens chip set. Image and carrier suppression of about 40 dB is also obtained.

The frequency synthesizer in these systems is off-chip. It typically consists of a single transistor oscillator, tuned by a ceramic resonator to meet the GSM specifications of very low noise sidebands, whose frequency is varied with a series or parallel varactor diode to synthesize the receive and transmit frequencies in a variable-modulus PLL. The entire synthesizer may drain 7 mA [32].

As the local oscillator in zero IF systems is tuned to the same frequency as the received carrier, it can radiate out of the antenna and interfere with nearby receivers tuned to slightly different frequencies. However, at the typical operating power levels and with the usual RF shielding in the handset, this is apparently not such a great problem that it requires special preventive measures [33].

ICs for Spread-Spectrum Wireless Transceivers

As cellular telephones saturate the available radio frequency allocations, spread-spectrum techniques are now being deployed in cordless telephones and wireless modems [34]. These usually operate in the two lower bands allocated by the FCC for the industrial, scientific, and medical (ISM) communities for unlicensed spread-spectrum communications: 902 to 928 MHz and 2.4 to 2.83 GHz. A ubiquitous wireless environment is envisioned, where mobile users, wherever they may be, may access data and communications services through an intricate network of base stations. The greatest hardware challenge is in developing the low-power, miniature handset.

A 700 kb/s frequency-hopped spread-spectrum transceiver by Plessey operating in the 2.4 GHz band, assembled on 2"×3" PCMCIA card for insertion into notebook computers [35], is a notable recent example of miniaturization. Data is transmitted by a binary frequency-shift keying of the carrier, and the carrier frequency is slowly hopped by a variable modulus PLL synthesizer to spread the spectrum across the entire band. All the active devices in the transceiver are on three ICs, consisting of a GaAs RF front-end, a silicon bipolar IF receiver, and a CMOS IC for the hopping frequency synthesis. Nevertheless, the transceiver requires 50 passive components, including six rather bulky filters, and it dissipates more than 1 W while transmitting 100 mW.

The single-chip GaAs IC front-end in this transceiver [36] is perhaps the most highly integrated of all the ICs discussed so far, as it includes the power amplifier and drivers for 2-GHz passive filters in the transmit and receive paths. It is otherwise a conventional double superheterodyne architecture. Interesting features are the DC series connection of the single-ended two-stage low-noise amplifier to reuse the same bias current, and that in its first version, the IC has more than 20 on-chip spiral inductors, perhaps the largest number found in any MMIC! In receive mode, the IC drains 30 mA from a 5 V supply, and with a on-chip RF switch, it can select one of two antennas to obtain spatial diversity.

ICs to Enable Future Transceivers

What will portable wireless communicators of the future look like? We may draw some conclusions from the foregoing summary of RF-IC developments to date. The crowded spectrum means that future wireless communicators will predominantly use spread-spectrum techniques, and the need for low-power dissipation will force a higher level of integration. Inspired by the modern paging receiver, which is perhaps the lowest energy wireless device in wide use, the author with his colleagues and their graduate students at UCLA is investigating the architecture and circuit design of a frequency-hopped, binary frequency-shift keyed, zero IF, all-CMOS two-chip transceiver capable of delivering up to 160 kb/s (the base ISDN rate) in the 900 MHz ISM band [37]. The transceiver ICs (Fig.7) will freely mix analog and digital circuits, and therefore CMOS is the IC technology of choice for the entire transceiver, *including* the RF front-end. Finally, these blocks will be implemented in an unmodified standard production CMOS process.

Low-power operation requires the entire system to operate on a 3V supply, determined mainly by the analog sections and the unmodified FET threshold voltages. The power otherwise wasted in driving pad and trace parasitic capacitances is avoided with the use of on-chip inductors, whose self-resonance after the process of selective removal of the silicon substrate (Fig.8) [38] is as high as it is on semi-insulating GaAs substrates. Thus, a 900 MHz RF amplifier with 30 dB gain draining only 3 mA has been built in 1-μm CMOS. The problems of the mixer in a direct conversion receiver, of attaining high linearity to suppress intermodulation with interferers, as well as potential leakage of the local oscillator signal through the antenna, are both circumvented with an unusual solution: a sub-sampling downconversion mixer (Fig.9) [39]. The circuit draws 4 mA to acquire samples of a 900 MHz modulated waveform at a 50 MHz rate which it directly translates to baseband. The linearity, as measured by a +22 dBm third-order intercept, exceeds that of most continuous-time 900 MHz monolithic mixers.

Key to a frequency-hopped transceiver is a low-power, agile frequency synthesizer with adequate spectral purity. A direct-digital frequency synthesizer (DDFS) [40] produces digital words representing samples of a sinewave at an arbitrary frequency set by an input control word. Followed by a suitably linear D/A Converter, this is the ideal means to implement an agile frequency source at baseband. A CMOS implementation of a DDFS-DAC at 3 V dissipates only 40 mW, and after upconversion mixing, it spreads the signal over 902-928 MHz (Fig. 10) [41]. The local oscillator consists of a four-stage MOS ring oscillator (Fig. 11) locked in a PLL to a lower frequency crystal reference. The quadrature outputs at 915 MHz for the image-reject mixers are tapped off at diametrically opposite points to a phase accuracy of a couple of degrees [42]. A binary-weighted array of FETs biased near threshold, followed by a matching network, may be digitally selected to deliver power levels as high as +15 dBm to the antenna with a 30% conversion efficiency.

If successful, such a "VLSI radio" would represent a major step forward in the gradual evolution of the integrated radio as described in this paper.

1. Hishida, Y. and M. Kazama, *Mobile Communication System: Its Semiconductor Technology,* in *Hitachi Review.* 1993, pp. 95-100.
2. Peil, W. and R.I. McFayden, "A Single Chip AM/FM Integrated Circuit Radio", *IEEE Trans. on Consumer Electronics,* v. CE-23(3), pp. 424-431, 1977.
3. Kanow, W. and I. Siewert, *Integrated Circuits for Hi-Fi Radios and Tuners,* in *Electronic Components and Applications.* 1981, pp. 11-27.
4. Okanobu, T., *et al.*, "A Complete Single-Chip AM/FM Radio Integrated Circuit", *IEEE Trans. on Consumer Electronics,* v. CE-28(3), pp. 393-408, 1982.
5. Kasperkovitz, W.G., "An Integrated FM Receiver", *Microelectronics Reliability,* v. 21(2), pp. 183-189, 1981.
6. Kasperkovitz, W.G., "FM Receivers for Mono and Stereo on a Single Chip", *Philips Technical Review,* v. 41(6), pp. 169-182, 1983.
7. van Dooremolen, W.H.A. and M. Hufschmidt, *A Complete FM Radio on a Chip,* in *Electronic Components and Applications.* 1983, pp. 159-170.
8. Okanobu, T., H. Tomiyama, and H. Arimoto, "Advanced Low-Voltage Single Chip Radio IC", *IEEE Trans. on Consumer Electronics,* v. 38(3), pp. 465-475, 1992.
9. Mitchell, D. and K.G. van Wynen, "A 150-mc Personal Radio Signaling System", *Bell System Technical J.,* v. XI(5), pp. 1239-1257, 1961.
10. Kerwien, A.E. and L.H. Steiff, "Design of a 150-megacycle Pocket Receiver for the BELLBOY Personal Signaling System", *Bell System Technical J.,* v. XLII(3), pp. 527-565, 1963.
11. Nagata, K., *et al.*, *Digital Display Radio Paging System,* in *NEC Research & Development.* 1983, pp. 16-23.
12. Vance, I.A.W., "Fully Integrated Radio Paging Receiver", *IEE Proceedings,* v. 129, Part F(1), pp. 2-6, 1982.
13. Holtvoeth, K., "A Fully-Integrated Low-Power FSK Receiver for VHF and UHF Paging", *IEEE Trans. on Consumer Electronics,* v. 35(3), pp. 707-713, 1989.
14. Wilson, I.F., *et al.*, "A Single-Chip VHF and UHF Receiver for Radio Paging", *IEEE J. of Solid State Circuits,* v. 26(12), pp. 1944-1950, 1991.
15. Burt, A., *Direct Conversion Receivers Come of Age in the Paging World,* in *GEC Review.* 1992, pp. 156-160.
16. Yamasaki, K., *et al.*, *Compact Size Numeric Display Pager with New Receiving System,* in *NEC Research & Development.* 1992, pp. 73-81.
17. Anderson, D. and R.J. Zavrel, *RF ICs for Portable Communications Equipment,* in *Electronic Components and Applications.* 1985, pp. 37-44.
18. Imai, Y., *et al.*, "Very Low-Current and Small Size GaAs MMICs for L-Band Front-End Applications", in *GaAs IC Symposium,* pp. 71-74, Monterey, CA, 1989.
19. MacCabe, S.P., "GaAs MMICs for Cellular Telephones, GPS and DBS Markets", in *WESCON Convention Rec.*, pp. 279-281, San Francisco, 1989.
20. Kermarrec, C., *et al.*, "High Performance, Low Cost GaAs MMICs for Personal Phone Applications", in *Int'l Symp. on GaAs and Related Compounds,* Karuizawa, Japan, 1992.
21. Sakai, H., *et al.*, "Low-Power GaAs ICs for Mobile Communication Equipment", in *Int'l Symp. on GaAs and Related Compounds,* Karuizawa, Japan, 1992.
22. Boveda, A., G.L. Bonato, and O. Ripolles, "GaAs Monolithic 1.5-2.5 GHz Image Rejection Receiver", in *Int'l Symp. on Circuits & Systems,* pp. 232-235, 1992.
23. Hart, P.J. and T.G.R. Hall, "A Low-Cost Approach to Cellular Mobile Telephones", *IEEE Trans. on Consumer Electronics,* v. 37(1), pp. 32-38, 1991.
24. Woo, C., *et al.*, "A Fully Integrated Transceiver Chip for the 900 MHz Communication Bands", in *GaAs IC Symposium,* pp. 143-146, 1992.

25. Tamura, Y., *et al.*, *Hand-Held Portable Equipment for Cellular Mobile Telephone*, in *NEC Research and Development*. 1987, pp. 34-43.

26. Tamura, Y., *et al.*, *Development of Advanced Mobile Telephone (Personal Pocket Phone)*, in *NEC Research and Development*. 1990, pp. 60-70.

27. Shimizu, I., *et al.*, "A New Pocket-size Cellular Telephone for NTT High-Capacity Land Mobile Communication System", in *IEEE Vehicular Technology Conference*, pp. 114-119, 1991.

28. Hikita, M., T. Tabuchi, and A. Sumioka, "Miniaturized SAW Devices for Radio Communication Transceivers", *IEEE Trans. on Vehicular Technology*, v. 38(1), pp. 2-8, 1989.

29. Suzuki, H. and T. Endo, *High-Density IC/LSI Package for Communication Electronics*, in *Hitachi Review*. 1993, pp. 141-144.

30. Fenk, J., *et al.*, "An RF Front-End for Digital Mobile Radio", in *Bipolar Circuits and Technology Meeting*, pp. 244-247, 1990.

31. Sevenhans, J., *et al.*, "An Integrated Si Bipolar RF Transceiver for a Zero IF 900 MHz GSM Digital Radio Frontend of a Hand Portable Phone", in *Custom IC Conference*, pp. 7.7/1-4, San Diego, CA, 1991.

32. Beyer, S. and G. Lipperer, *Low-Current Oscillator Design for 900 MHz GSM Applications*, in *Microwave Engineering Europe*. 1991, pp. 35-41.

33. Tsurumi, H. and T. Maeda, "Design Study on a Direct Conversion Receiver Front-End for 280 MHz, 900 MHz, and 2.6 GHz Band Radio Communication Systems", in *IEEE Vehicular Technology Conference*, pp. 457-462, St. Louis, MO, 1991.

34. Leonard, M., *Wireless Data Links Broaden LAN Options*, in *Electronic Design*. March 19, 1992, pp. 51-58.

35. Leonard, M., *PCMCIA-Sized Radio Links Portable WLAN Terminals*, in *Electronic Design*. August 5, 1993, pp. 45-50.

36. Devlin, L.M., *et al.*, "A 2.4 GHz Single Chip Transceiver", in *Microwave & Millimeter-Wave Monolithic Circuits Symp.*, pp. 23-26, Atlanta, GA, 1993.

37. Min, J., *et al.*, "A Low-Power Handheld Frequency-Hopped Spread Spectrum Transceiver Hardware Architecture", in *Virginia Tech's Third Symp. on Wireless Personal Communications*, pp. 10/1-8, Blacksburg, VA, 1993.

38. Chang, J.Y.-C., A.A. Abidi, and M. Gaitan, "Large Suspended Inductors on Silicon and their use in a 2-micron CMOS RF Amplifier", *IEEE Electron Device Letters*, v. 14(5), pp. 246-248, 1993.

39. Chan, P.Y., *et al.*, "A Highly Linear 1-GHz CMOS Downconversion Mixer", in *European Solid-State Circuits Conf.*, pp. 210-213, Sevilla, Spain, 1993.

40. Nicholas, H.T. and H. Samueli, "A 150-MHz Direct Digital Frequency Synthesizer in 1.25 micron CMOS with -90dBc Spurious Response", *IEEE J. of Solid State Circuits*, v. 26, pp. 1959-1969, 1991.

41. Chang, G., *et al.*, "A Low-Power CMOS Digitally Synthesized 0-13 MHz Agile Sinewave Generator", in *Int'l Solid State Circuits Conf.*, San Francisco, 1994.

42. Enam, S.K. and A.A. Abidi, "NMOS ICs for Clock and Data Regeneration in Gb/s Optical-Fiber Receivers", *IEEE J. of Solid State Circuits*, v. 27(12), pp. 1763-1774, 1992.

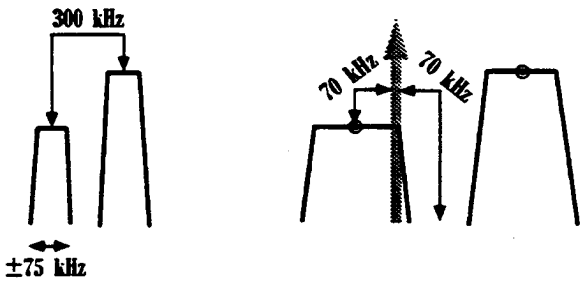


Fig.1. Channel spacing in the broadcast FM band, and the image frequency when the IF is reduced to 70 kHz [5,6].

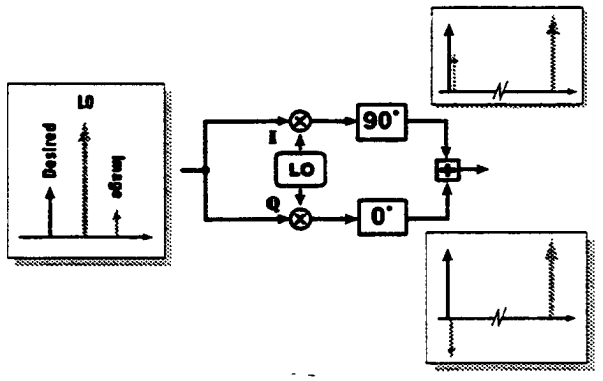


Fig.2. The image-cancelling mixer. This electronic alternative to a highly selective bandpass filter has only become practical on ICs, as the extent of image suppression depends on the matching of the two arms.

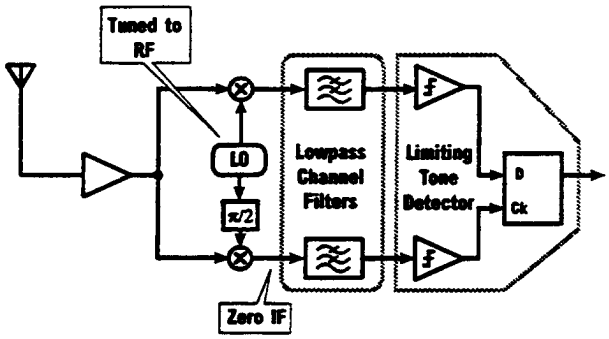


Fig.3. A simple zero IF receiver for binary FSK signals [12]. Including the lowpass channel filters,this may be readily integrated on a single IC.

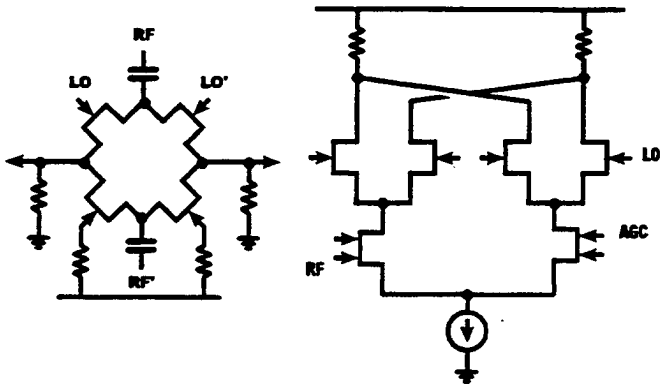


Fig.5. The double-balanced four FET switch mixer compared with a FET-tree mixer. The switch mixer affords higher linearity, but suffers from conversion loss. It is best suited for use in the transmitter upconverter.

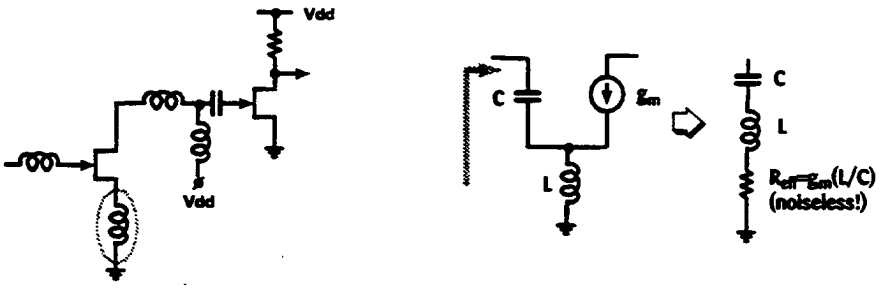


Fig.4. Typical GaAs low-noise RF amplifier uses inductors as matching elements, as tuning elements in the load, and as a noiseless degeneration element to present a resistive input impedance.

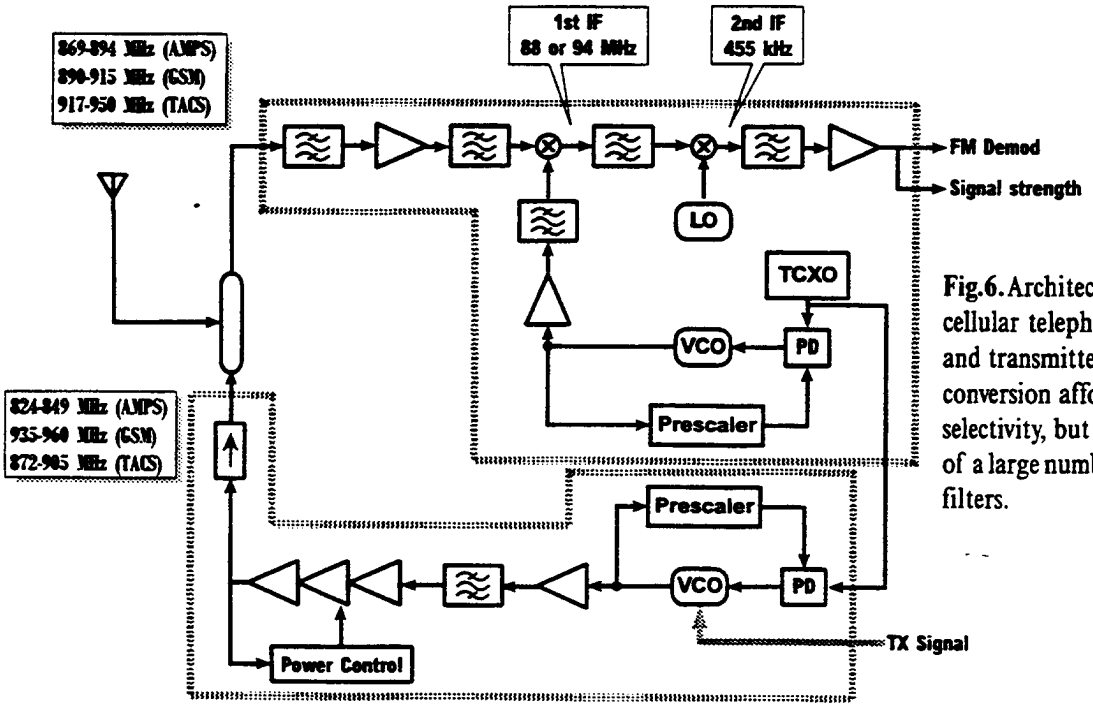


Fig.6. Architecture of a cellular telephone receiver and transmitter. Double conversion affords high selectivity, but at the expense of a large number of discrete filters.

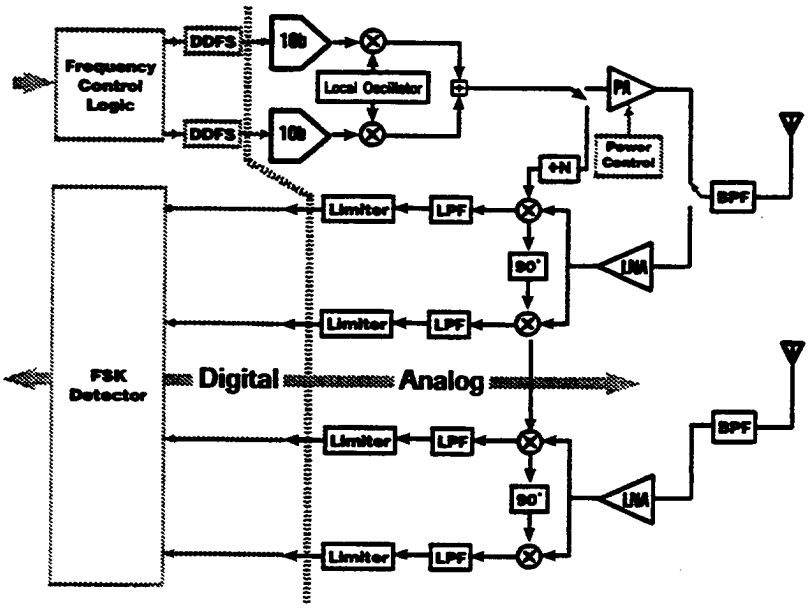


Fig.7. Architecture of the UCLA Frequency-Hopped FSK Transceiver. Two entire receive channels connected to the antennas help to combat multipath fading with spatial diversity.

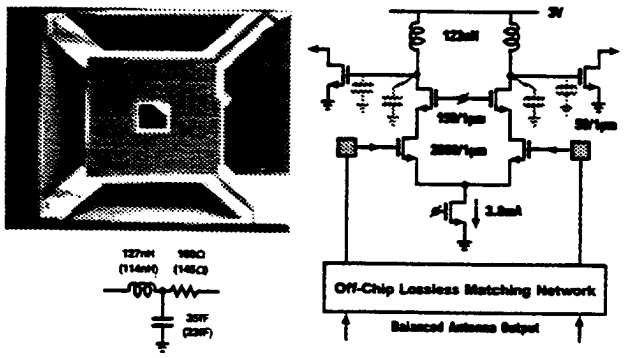


Fig.8. A 100 nH spiral inductor suspended over the silicon substrate. Self-resonance frequency is 3 GHz. A balanced CMOS RF amplifier affords 30 dB gain, 4 dB noise figure at 915 MHz.

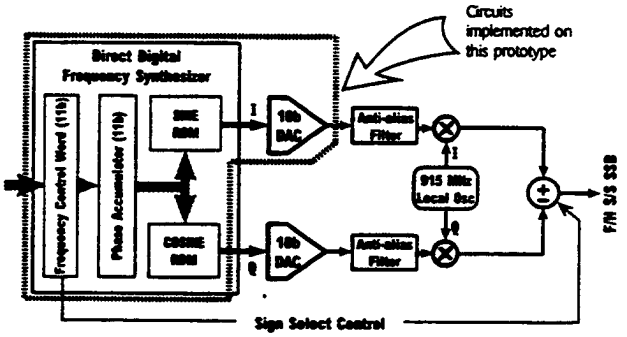


Fig.10. An agile frequency source using direct-digital frequency synthesis. Image reject mixer produces single-sideband, suppressed carrier output. Sign of summation at mixer outputs selects upper or lower sideband, so DDFS need only span 0 to 13 MHz to cover 902-928 MHz.

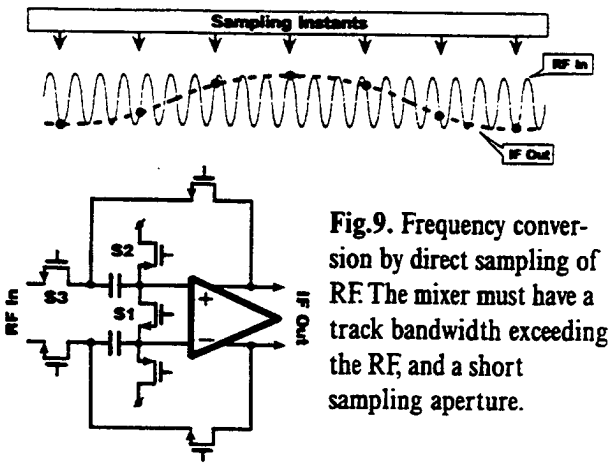


Fig.9. Frequency conversion by direct sampling of RF. The mixer must have a track bandwidth exceeding the RF, and a short sampling aperture.

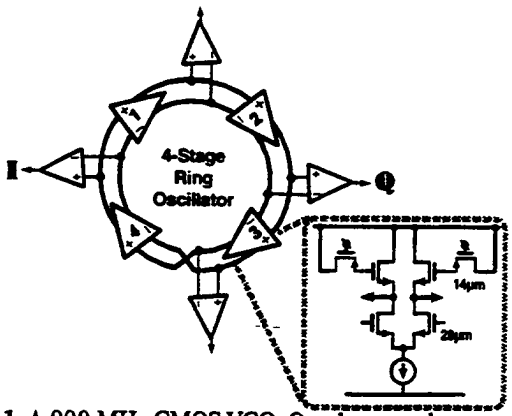


Fig.11. A 900 MHz CMOS VCO. Quadrature phases at any frequency tapped from diametrically opposite points.

An All-CMOS Architecture for a Low-Power Frequency-Hopped 900 MHz Spread Spectrum Transceiver

J. Min, A. Rofougaran, H. Samueli, and A. Abidi

Proceedings of the IEEE 1994 Custom Integrated Circuits Conference
pp. 379-382
May 1994

An All-CMOS Architecture for a Low-Power Frequency-Hopped 900 MHz Spread Spectrum Transceiver

Jonathan Min, Ahmadreza Rofougaran, Henry Samueli, and Asad A. Abidi

Integrated Circuits and Systems Laboratory
Electrical Engineering Department
University of California, Los Angeles

Abstract

An all-CMOS single chip architecture of a frequency-hopped spread spectrum (FH/SS) transceiver is presented. The transceiver implements complete modem functions spanning a tuned RF front-end to a digital binary frequency-shift keying (FSK) modulator/demodulator, all operating from a 3-V supply. This FH/SS transceiver is intended for use in a wide variety of indoor/outdoor portable wireless applications in the 902-928 MHz ISM band [1]. System features such as dual antenna/branch diversity, fast frequency hopping, and adaptive power control are incorporated into the transceiver architecture to achieve robust wireless digital data transmission over multipath fading channels. This paper describes the architectural features of the transceiver and surveys several of the key building blocks which have been completed to date.

I. Introduction

A FH/SS code division multiple access (CDMA) technique has advantages in that a) it provides an inherent immunity to multipath fading; and b) the signal processing is performed at the hopping rate, which is much slower than the chip rate encountered either in a direct sequence (DS) CDMA or time division multiple access (TDMA) system, thereby potentially resulting in much lower receiver power consumption. Furthermore, FSK modulation with noncoherent detection in a frequency-hopped system results in a much simpler transceiver architecture as compared with coherent amplitude and phase modulation methods commonly used in DS and TDMA systems. The proposed transceiver has the objective of passing data at rates up to 160 kb/s, hopping over the entire 26 MHz bandwidth. The FH/SS transceiver specifications are summarized in Table I.

The principal channel impairments in multi-user radio systems are attenuation due to shadowing, multipath fading, and interference from other radios. To overcome these impairments without resorting to high transmitter power, the proposed architecture incorporates many advanced techniques. Two antenna/receiver branches are used to provide equal-gain spatial diversity. Frequency hopping using a quadrature direct digital frequency synthesizer is implemented to provide frequency diversity. Adaptive power control is also employed with a programmable gain power amplifier so that the minimum transmitted power required for reliable communication is used, and channel coding using a convolutional code will be employed to improve the bit error rate performance.

TABLE I. FH/SS TRANSCEIVER SPECIFICATIONS

Multiple Access Scheme	Frequency-Hopped CDMA
Modulation Scheme	Noncoherent Binary FSK
Duplexing Mode	Time Division Duplex
Channel Coding	Rate 1/2, K=6 Convolutional Code
Hopping Bandwidth	26 MHz (902-928 MHz)
Data Rates	2 - 160 kbps
Maximum Hop Rate	160 khop/s
Antenna/Branch Diversity	2
Maximum Range	500 meters
Transmission Power	1 μW - 20 mW
Active Power	300 mW (Tx), 225 mW (Rx)
Operating Voltage	3 V
IC Technology	0.8 μm CMOS

Conventional RF front-end building blocks in the 1 GHz frequency band are typically implemented in GaAs or silicon bipolar technologies due to the required gain and low noise requirements. With rapid advances in CMOS IC technology, however, we have demonstrated a CMOS monolithic RF amplifier [2]. By using CMOS for the RF front-end, the entire transceiver is integrated in a single technology, with the exception of the antennas and 900 MHz passive bandpass filters. This full integration of RF, IF, and baseband components into one monolithic CMOS chip will result in substantial power savings. Further power reduction comes from architectural optimizations which will be described in the following section.

II. Transceiver Architecture

A. RF/IF Architecture

An RF/IF architecture employing high-performance low-

0-7803-1886-2/94 $3.00

power circuit techniques is proposed for implementing the FH/SS transceiver. The combination of a direct digital frequency synthesizer (DDFS) and a digital-to-analog converter (DAC) is used to generate a single-sideband 26 MHz spread-spectrum waveform in the 902-928 MHz band. The two DDFS/DAC channels produce sinewaves in quadrature with frequency selectable from 0 to 13 MHz (Fig. 1). After anti-alias filtering, these sinewaves are respectively upconverted by quadrature outputs from a 915 MHz local oscillator. If the two upconverted outputs are added, the output frequency ranges from 915 to 928 MHz; if they are subtracted, it ranges from 902 to 915 MHz. This signed I-Q frequency synthesis architecture reduces the highest frequency required from the DDFS/DAC to 13 MHz, yet covers the desired 26 MHz hopping bandwidth. In addition, this single-sideband (SSB) modulation technique eliminates the need for image rejection filters with sharp roll-offs, and therefore lowers the overall power consumption. By using DDFS techniques coupled with the inherent matching of monolithic CMOS analog circuits, we expect to achieve a SSB image rejection of 50 dB. Since we are using an FSK modulation scheme, a simple highly efficient class-C differential power amplifier is used for transmission. Finally, a low-cost off-the-shelf 902-928 MHz dielectric resonator bandpass filter is used between the power amplifier and the antenna to reject the out-of-band harmonics and meet FCC transmission mask requirements.

Before down conversion, the received signal needs to be amplified using a low noise amplifier (LNA). The gain and noise figure of the LNA are critical factors in determining the overall transceiver performance. We have successfully developed a new technique to achieve a monolithic high-Q inductor in CMOS without altering the fabrication process [2]. This technique is applied to design an LNA with high gain at 915 MHz. The high-Q inductor amplifier also provides extra filtering for rejecting the out-of-band signals and noise. To minimize the number of high frequency components and thereby save power, the dehopping and down conversion is performed in a single "direct-conversion" stage. Direct conversion requires accurate I-Q dehopping carriers at 915 MHz. Conventionally, I-Q carrier generation has been implemented by RC passive elements [3]. However, the accuracy of the I-Q signals is not adequate enough for high performance receivers. To overcome this short coming, a new technique is proposed which uses the I-Q outputs of the DDFS/DAC and of the LO. Accurate I-Q dehopping carriers are generated by taking the real and imaginary parts of the multiplications of these two complex signals, as shown in Fig. 1. Addition or subtraction will again result in selection of the upper side-band or lower side-band of the I-Q dehopping carriers. These carriers are then divided into lower frequencies to be used in the subsampling down conversion mixers. One drawback with this scheme, however, is that the circuitry which generates the I-Q dehopping carriers must be balanced. The input stage of the power amplifier is modified to add both of the I-Q signals.

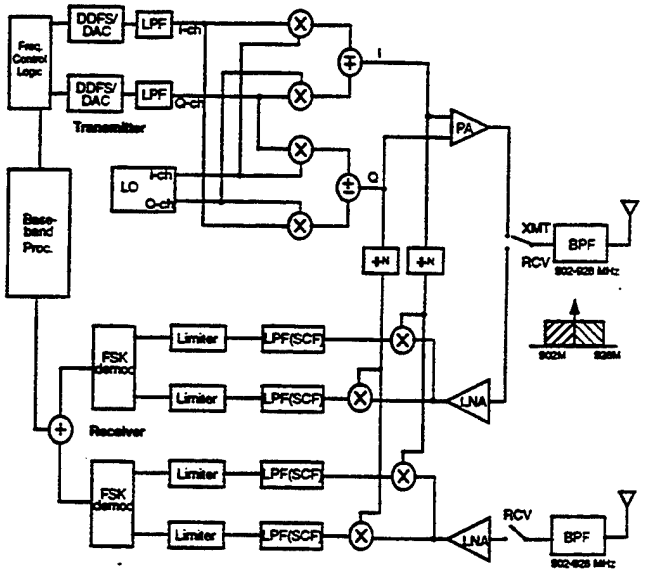


Fig. 1. FH/SS transceiver architecture.

This addition will result in another signal of the same frequency with a 45 degree phase shift only. Therefore, the transmitted signal carries the same information. The down conversion mixer must achieve a low intermodulation distortion (IMD). One way to reduce power and lower the IMD products is to use a sub-harmonic sampling circuit which is switched at an integer (small) fraction of the RF frequency (915 MHz). This method can be easily implemented using conventional switched-capacitor techniques which are well developed for CMOS technology. These techniques not only result in low IMD products but also result in lower overall power dissipation [4].

B. Baseband Architecture

Low power consumption requires that complex signal processing, such as adaptive equalization, be avoided if possible. At a given hop time, the FH system resembles a narrow band system unlike its DS counterpart. High rate frequency hoping with FSK modulation can be easily implemented by appropriately updating the frequency control word of the DDFS. The DDFS approach has several key advantages over conventional analog PLL-based synthesizers, the most important of which is being able to achieve rapid frequency changes with continuous phase transitions. The hop rate is no longer limited by the RF synthesizer settling time but by the system requirements.

Quadrature demodulation is required to efficiently use the entire 26 MHz transmission bandwidth. The automatic gain control (AGC) function is implemented using a combination of a lowpass filter (LPF), a hard limiter, and an oversampled 1-bit demodulator. The limiter works as a 1-bit quantizer. Due to odd harmonics produced by hard-limiting the sinewave, this approach limits the modulation scheme to binary FSK rather than M-ary FSK. However, the merits from power savings and hardware simplicity obtained from the absence of an

expensive (in both power and complexity) linear variable gain amplifier (VGA) and a complex analog-to-digital converter (ADC) justifies this architecture for a portable transceiver design. The quadrature binary FSK modulation scheme also relieves the problem of group delay variations at different frequencies in the LPF. The LPF is implemented using a switched-capacitor technique for its precise tuning capability. When programmable bandwidths and data rates are required, the SCF also has a distinct advantage over the continuous-time filter since its bandwidths scale with the clock rate. An all-digital quadrature demodulation architecture using an oversampled 1-bit correlation detector has been chosen over a fast fourier transform (FFT) or matched filter method. This choice is based on the fact that the correlator provides a flexible design which can easily accommodate programmable data rates, and the same hardware can be used to demodulate both data and sync hops. Following binary FSK demodulation, the rest of signal processing is all performed digitally. This includes frame synchronization, PN acquisition, clock recovery, and frequency tracking.

III. Analog and Digital Circuits

Various low-power analog and digital transceiver building blocks, covering RF, IF, and baseband circuits all developed in CMOS technology, are overviewed in this section. These prototype circuits will form the basis of the complete monolithic transceiver which is currently under development.

A. Low Noise RF Amplifier - A low-power, low-noise RF amplifier in 2-μm N-well CMOS has been implemented and tested [2]. The amplifier has been measured to provide 14 dB of gain while dissipating only 6.9 mW (excluding output buffers) with a single 3-V supply. The amplifier has been designed to be a tuned amplifier in order to limit the noise contribution from the amplifier itself. The tuning of the amplifier was achieved by a large on-chip spiral inductor resonating with the parasitic capacitances. A novel technique has been applied to suspend spiral inductors over air-filled pits in order to attain over 100 nH of inductance on chip while providing a high self-resonant frequency. The noise figure of this amplifier was measured to be 6.1 dB (excluding input terminations), and the center frequency was 770 MHz. A new version of this amplifier also has been developed in 1-μm CMOS which is centered at 915 MHz. This amplifier has a gain of 30 dB with a noise figure of 4 dB.

B. Power Amplifier - A low-power and highly efficient class-C power amplifier has been designed in a 1-μm N-well CMOS process for the 902-928 MHz application. The power amplifier transmits up to 20 mW with an efficiency of better than 35%. The maximum power consumption is about 60 mW. An on-chip inductor has been used to improve the gain and efficiency of the amplifier. A differential scheme is used to improve the performance of the power amplifier and minimize the out-of-band harmonics. This power amplifier uses a 6-bit binary word to control transmission power level. This technique provides more than 36 dB dynamic power control capability. To extend further, the input swing is also controlled with an additional 15 dB range, thus providing a total of 50 dB programmable dynamic range.

C. Sub-Sampling Mixer - A low-distortion sub-sampling mixer using 1-μm CMOS technology has been designed and fabricated [4]. This sub-sampling mixer is to be used in the receiver after the low noise RF amplifier stage to perform direct down conversion. The measured third-order intercept is about 25 dBm, and the fundamental output is compressed by 1 dB at an input power of 12 dBm. The overall power consumption is less than 11 mW. The clock of the sub-sampling mixer needs to run at a fraction the carrier frequency. In our measurements, a 50-MHz clock has been used to downconvert the 900 MHz signals.

D. Switched-Capacitor Lowpass Filter - A low-sensitivity 5th order elliptic switched-capacitor LPF has been designed. This filter rejects undesired high frequency images from the sub-sampler output by -60 dB. A continuous-time prototype based on the doubly-terminated LC ladder filter structure was derived first, and then by applying the bilinear transformation, a parasitic-insensitive switched-capacitor ladder filter was synthesized. The resulting circuit has very low sensitivities to both element-value variations and parasitic capacitances. High dynamic range and minimum capacitance spread were obtained through capacitance scaling. A 3-V prototype circuit fabricated in 1.0-μm CMOS technology dissipates about 15 mW. Further circuit optimization will be performed in the next version to reduce the power dissipation.

E. Local Oscillator (PLL/VCO) - A low phase noise local oscillator (LO) at 915 MHz is being developed. The LO must be designed to generate these high clock frequencies by scaling a stable reference source. The LO must not only have low phase/spurious noise but it must also have a low power dissipation. These specifications are further complicated by the large operating temperature range and low-power supply (3 V). A phase-locked loop (PLL) was chosen to keep the phase noise to a minimum. The individual circuit blocks will be optimized for low power dissipation. The PLL consists of an external crystal used to generate the reference frequency, a phase/frequency detector, a charge pump, a voltage controlled oscillator (VCO) and a digital divider. Simulated power dissipation is about 16 mW.

F. Digital-to-Analog Converter - A monolithic low-power 40 Msample/s 10-bit digital-to-analog converter (DAC) with a 3-V supply has been designed and tested [5]. Conventionally, high-speed and high-precision DACs are implemented with current mode architectures which typically dissipate a large power. The proposed DAC employs a pipelined charge redis-

tribution algorithm implemented differentially, in which the pipelined digital code controls MOS switches that charge the capacitors to some known reference voltage. Due to the inherent nature of a switched-capacitor circuit, there is no DC standing current running through the circuit itself. Thus, it is ideally suited for the low-power application. The core of the DAC circuit, excluding the digital clock generator and the output buffer, dissipates only 1 mW.

G. Direct Digital Frequency Synthesizer - A 3-V CMOS quadrature direct digital frequency synthesizer (QDDFS) has been fabricated [5]. The QDDFS synthesizes 10-bit output sine and cosine waves simultaneously at 40 Msample/s. Frequency is selected with an 11-bit word to obtain a 19.53 kHz resolution. The synthesizer can cover a bandwidth from DC to 20 MHz with a switching speed of 25 ns and a tuning latency of 2 clock cycles. Several techniques are employed to reduce the ROM storage. The worst-case spurious response for the proposed DDFS is -72.63 dB. The measured power dissipation of a 1.0-μm CMOS prototype at 40 MHz is about 25 mW.

H. FSK Demodulator - A low-power quadrature binary FSK correlation demodulator (Fig. 2) is in fabrication. Many architectural optimization techniques have been incorporated into the design. First, the front-end correlator uses only a 1-bit multiplier due to the fact that the incoming signal is hard limited to +/- 1's. Note that FSK demodulation requires only zero crossing information, not the amplitude. The input signal is oversampled according to the clock rate of a SCF/limiter combination (roughly 10 MHz). However, this is immediately decimated down to twice the baud rate following the integrate-and-dump block. The integration duration is half the baud time in order to generate the early and late signals for the clock recovery loop. A noncoherent (NC) demodulator requires a magnitude calculation unit, which is conventionally implemented with a pair of squaring multipliers followed by a summing node. In our architecture, an absolute value addition block replaces this squaring block. Thus, a truly multiplierless NC FSK demodulator has been implemented with little performance degradation. Simulated power dissipation in 1.0-μm CMOS technology is only 2 mW.

IV. Conclusions

In this paper, an efficient architecture for an all-CMOS FH/SS transceiver for use in low-power handheld communications applications has been presented, along with a description of the major analog and digital circuit components. Power savings of the transceiver comes from all three levels of optimization: system-level (FH over DS), architecture-level, and circuit-level (voltage scaling with parallelism and use of minimum geometry transistors when applicable). A highly integrated architecture of low voltage custom analog and digital CMOS components into a single chip, thereby minimizing the use of discrete components, is also a key to achieving the minimum power consumption.

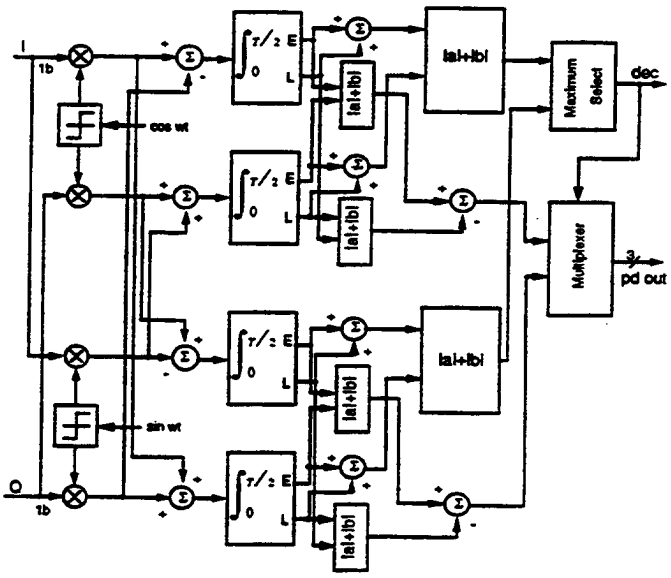


Fig. 2. Quadrature NC binary FSK demodulator.

Acknowledgment

The authors wish to thank all members of the FH/SS project team at UCLA for their technical contributions. This work was supported by ARPA/ESTO under Contract No. DAA B07-93C-C501 and the State of California MICRO Program.

References

[1] J. Min, A. Rofougaran, V. Lin, M. Jensen, H. Samueli, A. A. Abidi, G. Pottie, and Y. Rahmat-Samii, "A Low-Power Hanheld Frequency-Hopped Spread Spectrum Transceiver Hardware Architecture," *IEEE Workshop on VLSI in Communications*, Lake Tahoe, CA, Sep., 1993.

[2] J. Y-C. Chang, A. A. Abidi, and M. Gaitan, "Large Suspended Inductors on Silicon and Their Use in a 2-μm MOS RF Amplifier," *IEEE Electron Device Letters*, vol. 14, no. 5, pp. 246-248, May 1993.

[3] M. D. McDonald, "A 2.5 GHz BiCMOS Image-Reject Front-End," *Technical Digest of IEEE 1993 International Solid-State Circuits Conference*, San Francisco, CA, pp. 144-145.

[4] P. Y. Chan, A. Rofougaran, K. A. Ahmed, and A. A. Abidi, "A Highly Linear 1-GHz CMOS Downconversion Mixer," *Proceedings of 1993 European Solid-State Circuits Conference*, Sevilla, Spain, pp. 210-213.

[5] G. Chang, A. Rofougaran, M-K. Ku, A. A. Abidi, and H. Samueli, "A Low-Power CMOS Digitally Synthesized 0-13 MHz Sinewave Generator," to appear in *Technical Digest of IEEE 1994 International Solid-State Circuits Conference*, San Francisco, CA, Feb. 1994.

Characterisation of Electromagnetically Coupled Superquadric Loop Antennas for Mobile Communications Applications

M.A. Jensen and Y. Rahmat-Samii

IEEE Proceedings Microwaves, Antennas and Propagation
vol. 141, no. 2, pp. 85-93
april 1994

Characterisation of electromagnetically coupled superquadric loop antennas for mobile communications applications

M.A. Jensen
Y. Rahmat-Samii

Indexing terms: Antenna diversity, Loop antennas, Superquadrics

Abstract: Analysis of two mutually coupled loop antennas with arbitrary relative orientation and position and possibly different geometries is presented. The antennas are represented by generalised superquadric curves, which include circular, elliptical, and rectangular loop geometries, and they may be located either in a homogeneous region or next to an infinite ground plane. A Galerkin-type moment method with piecewise sinusoidal subsectional basis and weighting functions is used. Special consideration is given to implement the solution using curved wire segments instead of the commonly employed linear segments to improve computational efficiency. This very general computational tool is used to investigate the behaviour of coupled loops in configurations suitable for personal communications applications. A discussion of the use of antenna diversity to increase the received signal-to-noise ratio for communications equipment used in a multipath fading environment is also presented. Computational examples show that antenna diversity can provide significant improvements even for closely spaced loop antennas used in mobile communications applications.

1 Introduction

The increasing effort in miniaturisation of mobile communications equipment has spurred the development of small, low-profile antennas suitable for implementation in portable devices. In many instances, the circular or noncircular loop antenna provides an efficient radiator for the application. Whereas in the past a single antenna element has been used for mobile transceivers, the desire to combat multipath fading has led to the use of multiple elements arranged in a suitable diversity scheme. When more than one element is used, an important design consideration is the effect of mutual coupling on the antenna performance. The work presented here has been performed to provide an understanding of the behaviour of electrically coupled circular and noncircular loop radiators with arbitrary relative positions and orientations.

Paper 9904H (E11), first received 19th February and in revised form 10th August 1993

The authors are with the Department of Electrical Engineering, University of California, Los Angeles, 405 Hilgard Avenue, Los Angeles, CA, 90024-1594, USA

To our best knowledge, there has not been a unified solution methodology allowing analysis of very generally shaped loops with arbitrary orientations and positions. Past research efforts have to some extent addressed issues related to this work and have provided useful design data. For example, several theoretical investigations into the behaviour of circular loop antennas have been performed using a Fourier series expansion for the loop current distribution [1–7]. This analysis has been extended to the case of two identical, parallel loops arranged in either a coaxial arrangement [8] or with offset axes [9] to determine the mutual impedance between the elements. The characteristics of polygonal loops comprised of linear wire segments have been evaluated using a moment method solution approach [10]. The effects of relative antenna position and orientation on the mutual coupling between two closely spaced dipole antennas has also been demonstrated [11].

In this work, we present a unified formulation which can be used to characterise not only circular loops but also numerous other geometrical configurations, including square loops and folded dipoles, through use of the mathematical construction of the superquadric curve [12]. The methodology allows analysis of two coupled superquadric loops of possibly different geometries and with completely arbitrary relative orientation and position, located either in a homogeneous medium or near an infinite ground plane. The solution technique employs a piecewise sinusoidal Galerkin moment method [13] to determine the antenna current distribution from a coupled electric-field integral equation (EFIE) for thin wires [14]. By performing the integrations necessary for the moment method implementation along the curved path representing the antenna, the accuracy and efficiency of the formulation is enhanced [15]. The current obtained is subsequently used to find the self and mutual impedances, input impedance, directivity, and radiation pattern for the coupled-loop antennas. Both magnetic frill and delta gap source models are employed to allow investigation of different possible feeding scenarios.

Following presentation of the preliminary mathematical derivations associated with the superquadric curve geometry and moment-method solution, attention is turned to the concept of antenna diversity to mitigate the effects of multipath fading at a mobile terminal [16–19].

This work is funded by ARPA contract DAAB07-93-C-C501. M. Jensen's work is also supported under a National Science Foundation graduate fellowship.

Several key expressions are presented which provide a quantitative measure of the diversity performance of coupled antennas. Computational results for several representative isolated and coupled-loop configurations are then provided to illustrate the accuracy and flexibility of the solution methodology and to provide an understanding of the behaviour of loop antennas designed for wireless personal communications system applications.

2 Superquadric representation for loops

A superquadric curve [12] is the locus of points satisfying the equation

$$\left|\frac{x}{a}\right|^{v} + \left|\frac{y}{b}\right|^{v} = 1 \tag{1}$$

where a and b are the semiaxes in the x and y-directions, respectively, and v is a squareness parameter which controls the behaviour of the loop radius of curvature. A parametric representation for the superquadric curve is provided in Section 8. Fig. 1 illustrates the effect of the squareness parameter by presenting the curve geometry

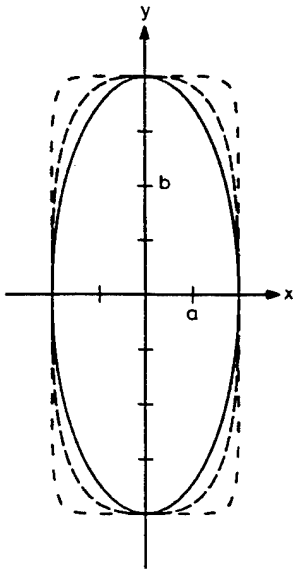


Fig. 1 *Superquadric geometry with aspect ratio of $b/a = 2$*
—— $v = 2$
- - - - $v = 3$
------- $v = 10$

for $v = 2$, 3, and 10 and an aspect ratio of $b/a = 2$. As can be seen from this Figure, a value of $v = 2$ corresponds to an ellipse while the loop squareness increases with v. It is evident that the superquadric representation allows modelling of numerous different loop configurations through variation of the shape parameters a, b, and v. Also, the model's capability to provide rounded corners on otherwise rectangular loops allows accurate representation of many practical geometries. This flexibility in geometrical configuration selection can be very useful for antenna packaging considerations.

Fig. 2 presents the geometry for the two coupled superquadric loops. Each loop is positioned in its own co-ordinate frame with the subscripts 1 and 2 denoting parameters associated with loops 1 and 2, respectively. The translation vectors r_{o1} and r_{o2} and the Eulerian angles [20] $(\alpha_1, \beta_1, \gamma_1)$ and $(\alpha_2, \beta_2, \gamma_2)$ designate the positions and orientations, respectively, of each loop co-ordinate frame with respect to the reference co-ordinate system. The co-ordinates s_1 and s_2 are the lengths of the curves measured from the feed points situated on each positive x-axis. The proper transformations required to

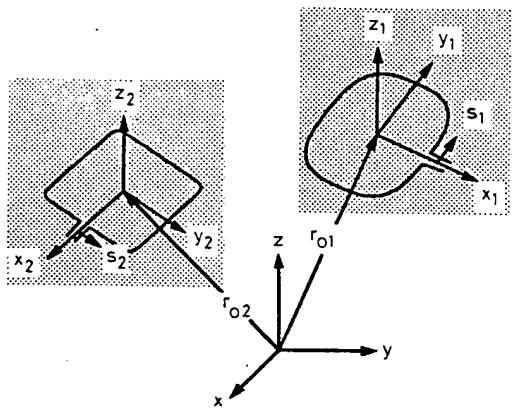


Fig. 2 *Geometry of two coupled loop antennas showing co-ordinates*

relate the loop co-ordinate systems to the reference frame in terms of the Eulerian angles and translation vectors are provided in Reference 20.

3 Moment method solution

3.1 Integral equation formulation

Analysis of the coupled superquadric loop antennas begins with the development of integral equations relating the currents on the wires to the excitation field applied at the antenna feed points. To formulate these integral equations, we used a thin-wire assumption which approximates the surface current density on the loops as filamentary currents $I_1(s_1)$ and $I_2(s_2)$ which flow in the axial directions only. The thin-wire form of the electric-field integral equation [14] can be manipulated to incorporate the effects of mutual coupling between the two loops as in

$$\begin{aligned}
&\hat{s}_1 E^{inc}(s_1) \\
&\quad = \frac{-1}{4\pi j\omega\varepsilon_0}\left\{\int_{L_1}\left[k_0^2\,\hat{s}_1\hat{s}'_1 I_1(s'_1)G_{11} + \frac{\partial I_1(s'_1)}{\partial s'_1}\frac{\partial G_{11}}{\partial s_1}\right]ds'_1\right. \\
&\qquad \left. + \int_{L_2}\left[k_0^2\,\hat{s}_1\hat{s}'_2 I_2(s'_2)G_{12} + \frac{\partial I_2(s'_2)}{\partial s'_2}\frac{\partial G_{12}}{\partial s_1}\right]ds'_2\right\} \\
&\hat{s}_2 E^{inc}(s_2) \\
&\quad = \frac{-1}{4\pi j\omega\varepsilon_0}\left\{\int_{L_1}\left[k_0^2\,\hat{s}_2\hat{s}'_1 I_1(s'_1)G_{21} + \frac{\partial I_1(s'_1)}{\partial s'_1}\frac{\partial G_{21}}{\partial s_2}\right]ds'_1\right. \\
&\qquad \left. + \int_{L_2}\left[k_0^2\,\hat{s}_2\hat{s}'_2 I_2(s'_2)G_{22} + \frac{\partial I_2(s'_2)}{\partial s'_2}\frac{\partial G_{22}}{\partial s_2}\right]ds'_2\right\}
\end{aligned} \tag{2}$$

where

$$G_{pq} = \frac{e^{-jk_0R_{pq}}}{R_{pq}}$$

and the unit vectors $\hat{s}_1$ and $\hat{s}_2$ are in the direction of the co-ordinates s_1 and s_2, respectively. L_1 and L_2 denote integration over each loop perimeter. The constant $k_0 = 2\pi/\lambda$ is the free-space wavenumber and R_{pq} represents the distance from the observation point on loop p to the source point on loop q, where $p, q = 1, 2$. Although the integrations in eqn. 2 are performed in the local co-ordinates of each antenna, the components of the unit and field vectors must be expressed in the reference co-ordinate system to ensure proper evaluation of the vector inner products. These integral equations can be aug-

mented to allow analysis of loops over an infinite ground plane through addition of terms representing the loop image.

When computing the value of R_{pq}, the thin-wire assumption allows the approximation that the observation point is along the wire axis rather than the wire surface. The source point is chosen according to the convention that the point is on the wire surface at $z_p = r_{wp}$ if $p = q$ (r_{wp} = wire radius of loop p), and on the wire axis if $p \neq q$. Use of this convention provides a convenient method for overcoming the difficulties associated with the singular integrand for coincident source and observation points. Computation of the distance R_{pq} is most conveniently performed in the reference co-ordinate system.

3.2 Moment method solution: curved subsections

To apply the integral equations of eqn. 2 to the superquadric geometry, we used the parametric representation developed in Section 8. This involves changing the functions of s to functions of the parameter τ and integrating over the range $0 \leqslant \tau \leqslant 2\pi$. An important feature of the use of this parametric construction for the superquadric loop is that curved wire segments, rather than the commonly used piecewise-linear segments, are implemented to represent the geometry. Recent developments have shown that such a practice proves to be more computationally efficient [15].

The moment method matrix is formed from the integral equations by dividing the curve into discrete segments at the points τ_n, $n = 1$ to N. A Galerkin moment method approach with piecewise sinusoidal subsectional basis and testing functions of the form

$$f_n(\tau) = \begin{cases} \dfrac{\sin(\tau - \tau_{n-1})}{\sin(\tau_n - \tau_{n-1})} & \text{if } \tau_{n-1} \leqslant \tau \leqslant \tau_n \\ \dfrac{\sin(\tau_{n+1} - \tau)}{\sin(\tau_{n+1} - \tau_n)} & \text{if } \tau_n \leqslant \tau \leqslant \tau_{n+1} \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

is then implemented to create the linear system whose unknown vector represents the currents $I_1(s_1)$ and $I_2(s_2)$ on the coupled loop antennas.

The excitation vector in the matrix equation is obtained using either the delta gap or magnetic frill source model. In the delta gap model, the impressed field is assumed constant and to exist only in a small feeding gap in the loop conductor. The field value is given by $\hat{s} \cdot \boldsymbol{E}^{inc} = V/\Delta g$, where Δg is the gap width and V is the impressed feed voltage. In general, Δg is not constrained to equal the subsectional arc length. The magnetic frill source model [6, 21] uses an annular ring of magnetic current at the antenna feed point to represent the excitation from a coaxial line feeding a half loop over a large ground plane. Using image theory, the half loop with the ground plane can be represented as a full loop with the magnetic frill plus its image, as shown in Fig. 3. The inner and outer radii of the annular current ring correspond to the coaxial inner and outer wire radii r_w and a_0, respectively, which are related by $a_0 = 2.3r_w$ for an air-filled 50 Ω coaxial cable. It is important to recognise that the current and input admittance of the half loop are twice those of the full loop antenna model of Fig. 3.

Once the currents $I_1(s_1)$ and $I_2(s_2)$ have been determined from the coupled integral equations, the far-field radiation pattern for the superquadric loops can be obtained from the radiation integral. As shown later, evaluation of the antenna diversity performance of the coupled antennas requires distinction between the contribution from each loop to the overall pattern. Therefore,

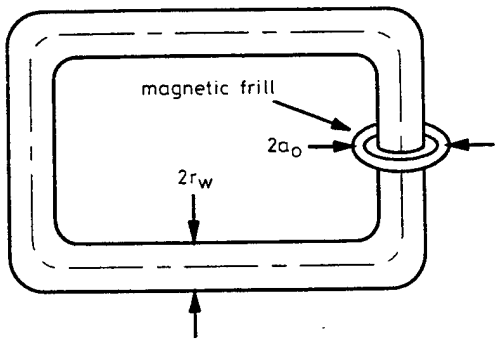


Fig. 3 *Full loop model of half-loop excited by magnetic frill source*

we express the pattern due to each antenna as

$$\boldsymbol{E}_p(\theta, \phi) \simeq \frac{-j\omega\mu_0}{4\pi} \frac{e^{-jk_0 r}}{r} \int_{L_p} I_p(s'_p)\hat{s}'_p \exp(jk_0 \hat{\boldsymbol{r}} \cdot \boldsymbol{r}'_p)\, ds'_p \quad (4)$$

for $p = 1, 2$, where r is the distance from the co-ordinate origin to the field observation distance and $\boldsymbol{r}'_p$ is the position vector as defined in Fig. 2. Once again, the vector components in eqn. 4 should be expressed in the reference co-ordinates to allow proper evaluation of the vector inner products. The total pattern may then be written as the sum

$$\boldsymbol{E}(\theta, \phi) = \boldsymbol{E}_1(\theta, \phi) + \boldsymbol{E}_2(\theta, \phi) \quad (5)$$

4 Mobile communications scenario

One of the objectives of this work is to determine the performance of the coupled superquadric loop antennas when used in a diversity scheme for a mobile communications system. In this application, we are interested in the use of space, angle, and polarisation diversity to combat the effects of short-term or Rayleigh-type fading in a multipath environment. This type of fading occurs when multiple waves with random phases exist simultaneously to produce a spatial interference pattern, causing the received signal strength to vary with antenna position. Antenna diversity operates on the concept that by providing multiple antennas for a single receiver, there is an increased probability that at least one of the antenna elements will receive a signal of adequate strength. Such antenna diversity scenarios are becoming more predominant as communications systems demand increased signal quality and reliability.

Quantitative evaluation of the performance of multiple antennas configured in a diversity arrangement involves investigation of the statistics of the signals received by the elements through examination of the envelope correlation coefficient ρ_e. This quantity provides a measure of the similarity of the voltages at each antenna terminal and should ideally be zero for maximal diversity performance (assuming Rayleigh distributed field strengths). For simplicity, the discussion here is concerned with two antennas only, although additional antennas may be analysed by the same procedure [16–19].

Although the statistical behaviour of the fields in a multipath fading environment as well as the expressions for the correlation coefficient for two antennas have been addressed at length in the literature [16–17], a brief presentation of the key assumptions made in developing these expressions is provided. In all expressions, the variables θ and ϕ are taken with respect to a co-ordinate system

oriented with the z-axis perpendicular to the earth. The following are the basic assumptions applied in the theoretical analysis:

(*a*) The Rayleigh probability density function describes the envelope of the fading signal, leading to zero-mean complex Gaussian descriptions of the voltages received by the two antennas [17].

(*b*) Orthogonal polarisations in the incoming multipath wave are uncorrelated, equally likely, and characterised by power densities per steradian $S_\theta(\theta, \phi)$ and $S_\phi(\theta, \phi)$ for the $\hat{\theta}$ and $\hat{\phi}$ polarisations, respectively.

(*c*) Each individual polarisation is spatially uncorrelated (i.e. wave incoming at (θ, ϕ) is uncorrelated with wave at (θ', ϕ') for $\theta' \neq \theta$ or $\phi' \neq \phi$).

(*d*) The spatial distribution of the incoming multipath waves is limited to the horizontal plane only ($\theta = \pi/2$) and the power density for each wave is a constant over this plane (i.e. $S_\theta(\pi/2, \phi) = S_\theta^0 =$ constant and $S_\phi(\pi/2, \phi) = S_\phi^0 =$ constant).

Using these assumptions leads to an expression for the envelope correlation coefficient for the signals received by the two antennas which assumes the form [16]

$$\rho_e = \frac{\left| \int_0^{2\pi} \boldsymbol{E}_1(\pi/2, \phi) \cdot \boldsymbol{E}_2^*(\pi/2, \phi)\, d\phi \right|^2}{\int_0^{2\pi} |\boldsymbol{E}_1(\pi/2, \phi)|^2\, d\phi \int_0^{2\pi} |\boldsymbol{E}_2(\pi/2, \phi)|^2\, d\phi} \tag{6}$$

where $\boldsymbol{E}_p(\theta, \phi) = E_{\theta p}(\theta, \phi)\hat{\theta} + E_{\phi p}(\theta, \phi)\hat{\phi}$ is the vector radiation pattern associated with antenna p obtained from eqn. 4 for the case of the coupled loop antennas.

There arises some question concerning precisely which currents to use to obtain the patterns $\boldsymbol{E}_p$ in eqn. 6. For instance, isolated or coupled loop configurations may be used in the computation for the current on each loop. If a coupled loop configuration is used, the second antenna may be either excited or terminated in an open circuit or matched load. In the examples shown in this work, the currents on each antenna are computed by using identical excitations for both antennas. This scheme fully includes the effects of mutual coupling on the antenna current behaviour. The currents obtained are subsequently placed in eqn. 4 to compute the patterns associated with each loop.

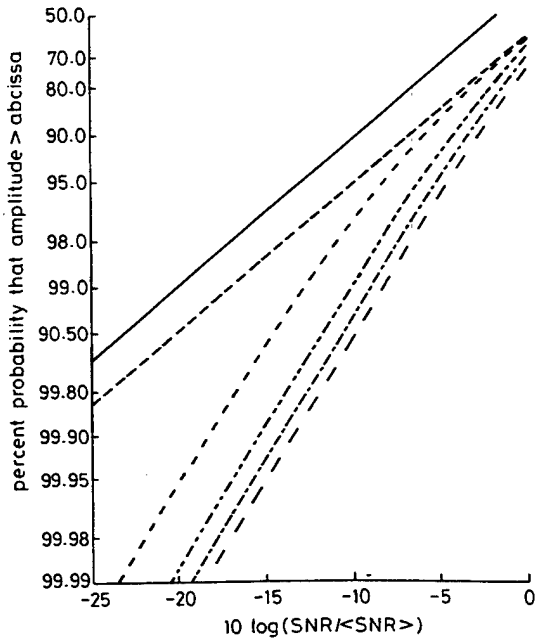


Fig. 4 *Probability distribution of received SNR for two-branch maximal ratio combining for different values of envelope correlation coefficient* ρ_e

——— Rayleigh ------- $\rho_e = 0.9$ —·—· $\rho_e = 0.3$
– – – – $\rho_e = 1.0$ —··— $\rho_e = 0.6$ — — $\rho_e = 0.0$

Fig. 4 shows the cumulative probability distribution of the signal-to-noise ratio (SNR) normalised to its average ($\langle$SNR$\rangle$) for various values of ρ_e and for two-antenna diversity. Maximal ratio combining in which the two signals are received, cophased, properly weighted, and added is assumed in this plot. Also shown is the Rayleigh distribution which corresponds to a single antenna in the multipath environment. As can be seen, reduction of the envelope correlation coefficient provides a considerable increase in the probability of receiving a signal of adequate strength for reliable communication. In light of this, the goal in diversity antenna design is to minimise this factor to the extent possible using a combination of spatial, angle, and polarisation diversity.

5 Computational examples

The preceding developments can be used to perform some computational studies on the characteristics of coupled superquadric loop antennas. Throughout this Section a wire radius r_w is used such that

$$\Omega_0 = 2 \ln \frac{P}{r_w} = 10 \tag{7}$$

is satisfied, where P is the loop perimeter. Five-point Gaussian quadrature integration is used to evaluate the integrals over each subsection. Where the magnetic frill generator source is implemented, the frill dimensions are chosen to match those of a 50 Ω feeding coaxial line. Numerical tests of the moment method algorithm were performed to ensure that small enough subsections were used to obtain convergent values for the input impedance.

5.1 Parallel circular loops

Before demonstrating the versatility of the computational model, it is interesting to investigate the correspondence between the results of this work and previously obtained results for parallel circular coupled loop antennas. Fig. 5 illustrates a typical plot of self and mutual admittance against separation z_0 for the case of two circular loops with circumferences P of 0.2λ, 0.4λ, 0.6λ, and 0.8λ. The magnetic frill source was used for the configuration shown in the inset of Fig. 5*a*. This plot indicates the increase in mutual coupling with increased loop dimensions characterised by higher mutual admittance values and more pronounced admittance variation with separation distance. Also noteworthy is the reduction in mutual admittance with increased separation, resulting in self admittances which asymptotically approach the input admittance values for isolated loops of $0.00603 - j3.36$, $0.0370 - j0.494$, $0.173 + j1.28$, and $0.900 + j3.38$ mS for loop sizes of 0.2λ, 0.4λ, 0.6λ, and 0.8λ, respectively. Finally, this computation illustrates the low admittance (high impedance) values which result in the impedance matching problem for loops near 0.5λ. The dots in the Figure are values taken from a study by Iizuka *et al.* [8]. It is clear that excellent agreement is obtained between the two sets of results over a wide range of loop sizes and separations. This comparison supports the model's capability to faithfully reproduce the results of previous studies and provides insight into the effect of antenna separation on the mutual admittance of the coupled loops.

5.2 Current distribution

When two superquadric loop antennas are operating in close proximity to each other, it is expected that the

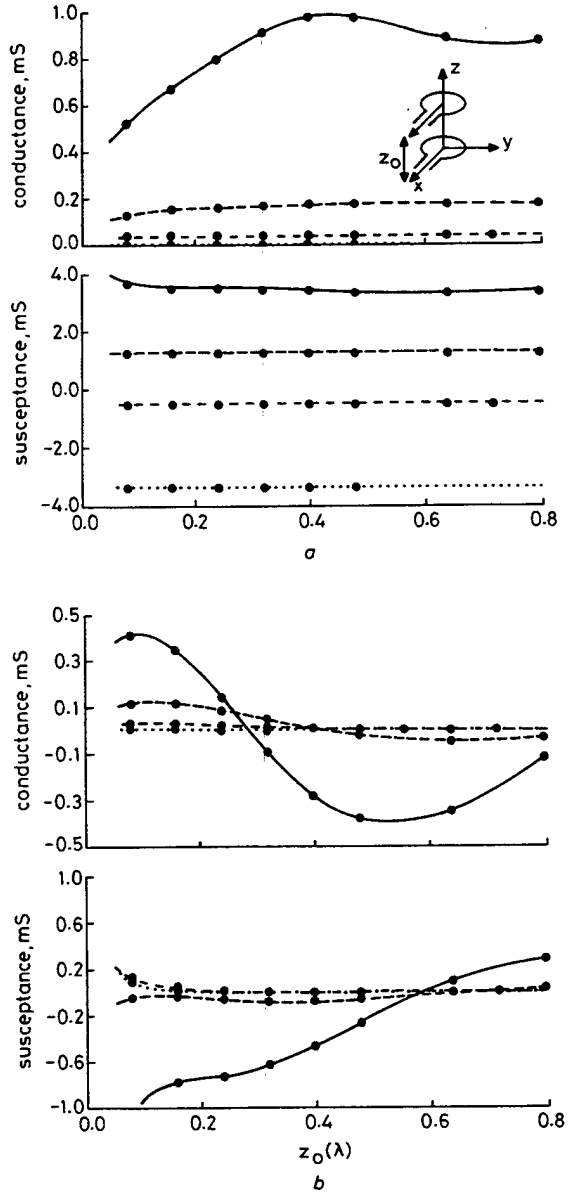


Fig. 5 *Self admittance against separation distance for parallel circular loops with different loop circumferences P*

a Self admittance
b Mutual admittance
——— $P = 0.8\lambda$
– – – – $P = 0.6\lambda$
------- $P = 0.4\lambda$
····· $P = 0.2\lambda$
●●● Reference 8

mutual coupling will affect the current distribution along the wires. Fig. 6 illustrates this effect for a nearly rectangular 0.8λ loop with $v = 30$ and $b/a = 2$. The current magnitude is normalised to the antenna input voltage. In this example, the solid line represents the current magnitude and phase for the loop when it operates in phase with an identical loop located $z_0 = 0.2\lambda$ from and parallel to the antenna. The dashed line represents the current distribution for the same loop operating in an isolated environment. As can be seen, although the general behaviour of the current is unchanged, the levels and detailed variation of the current are noticeably influenced by the presence of the second antenna. Especially noteworthy is the substantial decrease in the current magnitude along the wire near the point opposite the feed at $s = 0.4\lambda$.

5.3 Loops near an infinite ground plane

One of the practical uses of the loop is in communications equipment, where the loop may possibly be near a

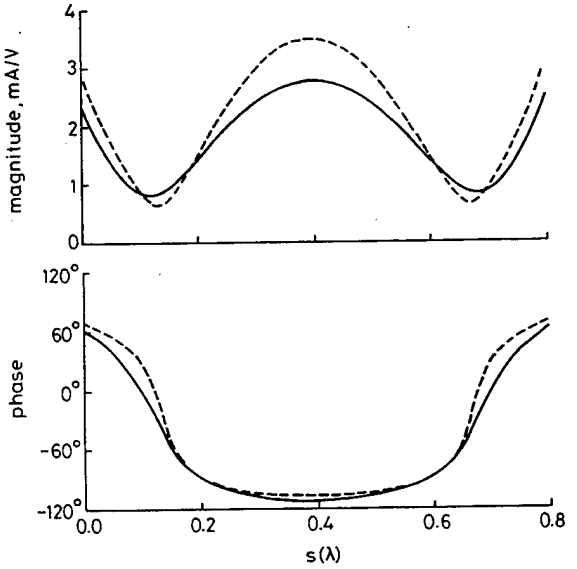


Fig. 6 *Current distribution against perimeter co-ordinate s for a 0.8λ loop ($v = 30$, $b/a = 2$) parallel to and located $z_0 = 0.2\lambda$ from identically shaped and excited antenna compared to that of isolated loop*

——— coupled loops
– – – – isolated loop

finite-size conducting plate or other body. A first-order approximation of the behaviour of the loop in this environment may be obtained by solving the problem for the loop near an infinite ground plane. For example, Fig. 7*a* shows a 0.2λ nearly rectangular ($v = 50$) loop located $c = 0.01\lambda$ from an infinite planar conductor for three different positions of the source. The variation of the input impedance versus the aspect ratio b/a is illustrated in Fig. 7*b* for a delta gap generator with $\Delta g = 0.003\lambda$. It is evident that both the source location and the aspect ratio play a role in determining the loop input impedance for this configuration.

Fig. 8 shows the far-field radiation patterns obtained for a 0.25λ elliptical loop ($v = 2$) of aspect ratio $b/a = 2$ oriented parallel to and located 0.1λ from the ground plane as illustrated in the inset. For comparison, the pattern for the same loop isolated from the ground plane is also provided. The patterns are normalised to represent the antenna directivity and are shown for the principal planes only. As is expected, the presence of the conducting sheet results in a null field at $\theta = 90°$ and noticeably increases the boresight antenna directivity. The E_θ component in the $\phi = 0°$ cut is not included because it experiences a null in this plane. Table 1 compares the input impedance and total real power radiated by the loop in Fig. 8 for the isolated and ground plane configurations. As expected, the presence of the ground plane significantly reduces the total power radiated by the loop antenna.

Table 1: Input impedance and radiated power (normalised to square of input voltage) comparison for loop configurations of Fig. 8

Configuration	Input impedance	Radiated power
		mW/V²
Ground plane	$0.267 + j383.94\ \Omega$	0.0543
Isolated	$1.382 + j385.65\ \Omega$	0.279

5.4 Concentric loops

In many systems, multiple loop antennas may be positioned such that they are in a concentric configuration.

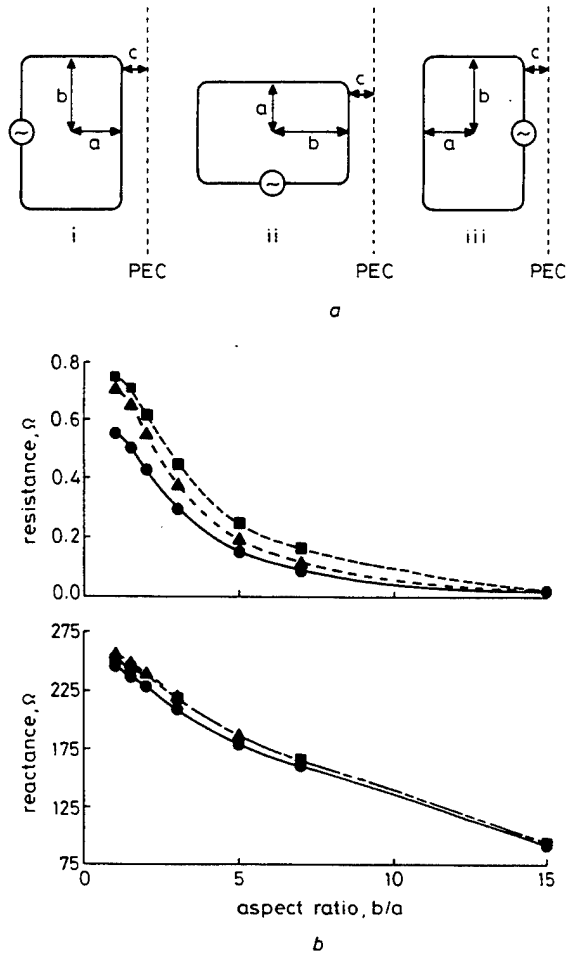


Fig. 7 *0.2λ loop with v = 50 located c = 0.01λ from infinite conducting plane for three source positions*

a Geometry
b Input impedance against aspect ratio *b/a* for three cases shown
●—● case i
■—■ case ii
▲—▲ case iii

An example of this type of situation is illustrated in the inset of Fig. 9, where the larger and smaller loops each have a 0.25λ perimeter at their operating frequencies of f_0 and $3f_0$, respectively. The loops have aspect ratios of $b/a = 1$ and are excited using a delta-gap source model. For these computations, each loop is in turn excited at its operating frequency while a short circuit is placed across the terminals of the other loop. The curves in Fig. 9 compare the variation in the input impedance versus the squareness parameter v for each loop as well as that of an isolated 0.25λ loop. From these results, we conclude that the presence of the second loop has a significant effect on the impedance values. This plot further illustrates the role of the loop squareness in determining its terminal impedance. These results can be very useful in the design of coupled loop antennas for communications applications.

A second interesting configuration is shown in the inset of Fig. 10 where two loops are arranged in a cross geometry and placed at $z_0 = 0.26\lambda$ above an infinite ground plane. Each loop is fed using a delta gap source model and has superquadric parameters of $v = 10$ and $a = 2.5b = 0.125\lambda$. Fig. 10*a* and *b* provide the far-zone radiation patterns in dB normalised to represent the antenna directivity in the *xz* (or *yz*) and *xy*-planes, respectively, when the two loops are fed in phase. Fig. 10*c* and *d* provide the directivity patterns in dB in the *xy*-plane for the same configuration when the two loops are

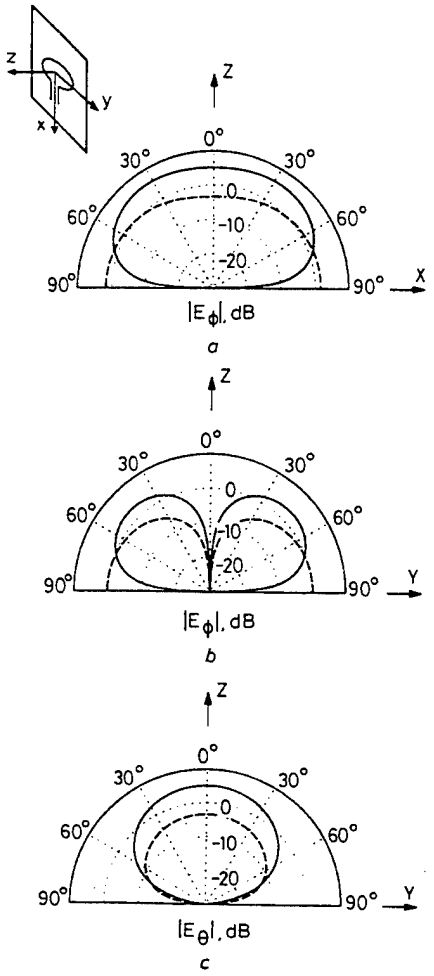


Fig. 8 *Directivity pattern for 0.25λ elliptical loop of aspect ratio b/a = 2 both isolated and located 0.1λ from infinite ground plane*

a $\hat{\phi}$-polarisation, $\phi = 0°$ plane (E_θ is null)
b $\hat{\phi}$-polarisation, $\phi = 90°$ plane
c $\hat{\theta}$-polarisation, $\phi = 90°$ plane
——— ground plane
- - - - isolated

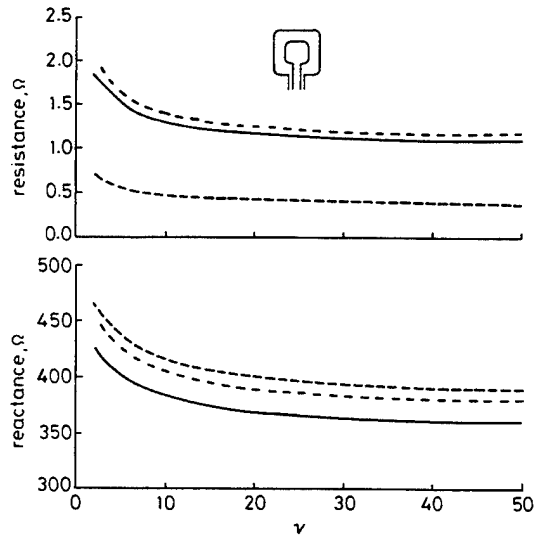


Fig. 9 *Input impedance against squareness parameter v for concentric 0.25λ loops with b/a = 1 compared with that of isolated loop*

Larger and smaller loops operate at f_0 and $3f_0$, respectively
——— single loop
- - - - small loop
------- large loop

fed 45° and 90° out of phase, respectively. As can be seen, through proper variation of the feeding voltages, the antenna radiation pattern can be made to be omnidirectional in the horizontal plane.

Fig. 12 illustrates the effects of antenna rotation on the envelope correlation coefficient for the 0.25λ loop with $v = 10$ and $b/a = 1$. Results for antenna separations y_0 of 0.1λ, 0.2λ, 0.3λ, and 0.4λ are given as a function of the

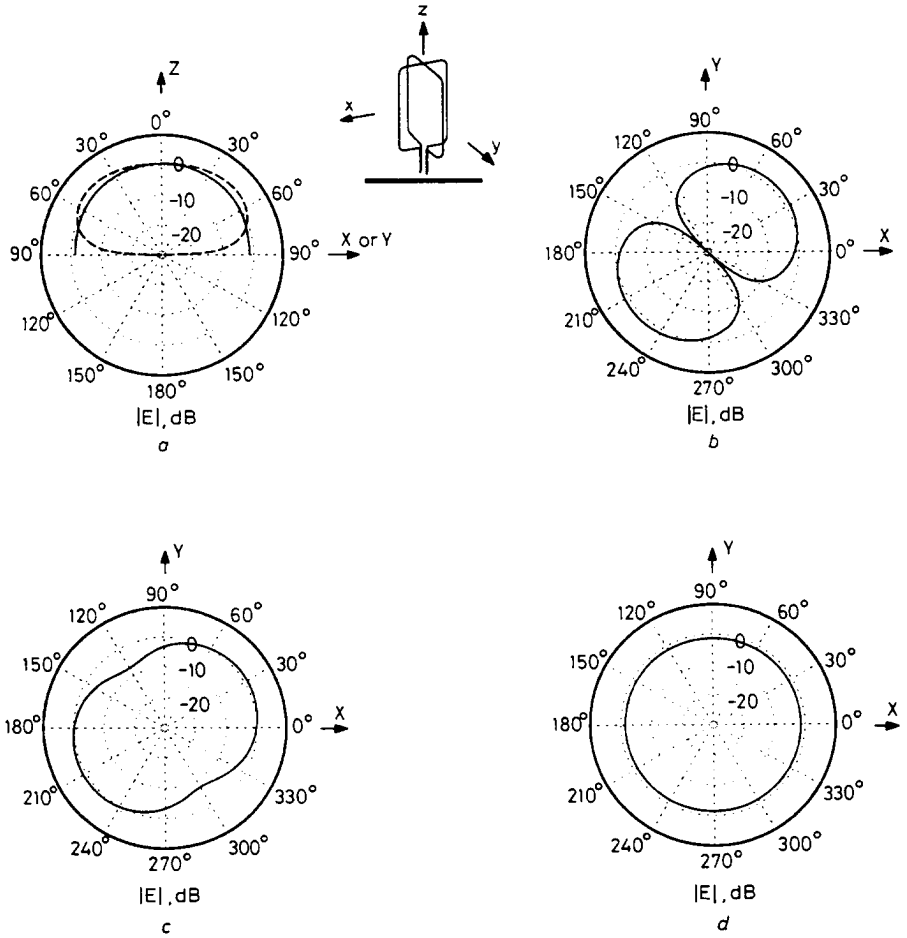


Fig. 10 *Directivity patterns in dB for two crossed superquadric loops at $z_0 = 0.26\lambda$ over infinite ground plane with $v = 10$ and $a = 2.5b = 0.125\lambda$*

a *xz*-plane pattern with loops excited in phase
b *xy*-plane pattern with loops excited in phase
c *xy*-plane pattern with loops excited 45° degrees out of phase
d *xy*-plane pattern with loops excited 90° degrees out of phase
——— E_θ
- - - - E_ϕ

5.5 Diversity

The diversity performance of small antennas is an important consideration in the design of personal communications networks. To determine the diversity performance of a given antenna configuration, the loops are each excited with the same feed voltage and the currents are determined using the formulation presented in this paper. The envelope correlation coefficient is then computed using eqn. 6 in conjunction with eqn. 4. Fig. 11 provides a plot of the envelope correlation coefficient as a function of loop separation y_0 for the case of two parallel loops as depicted in the inset. In this Figure, the solid line represents ρ_e for two antennas with isotropic patterns in the xy-plane [19]. The remaining curves show ρ_e for loops with perimeters 0.25λ, 0.50λ, 0.75λ, and 1.0λ with $v = 10$ and $b/a = 1$. This plot demonstrates the well known theortical zero in the correlation coefficient for antenna spacings near 0.4λ, which would therefore be the ideal operating point. However, in smaller systems where this separation is not physically possible, reasonable diversity performance can still be achieved for spacings as low as 0.15λ where $\rho_e \simeq 0.7$.

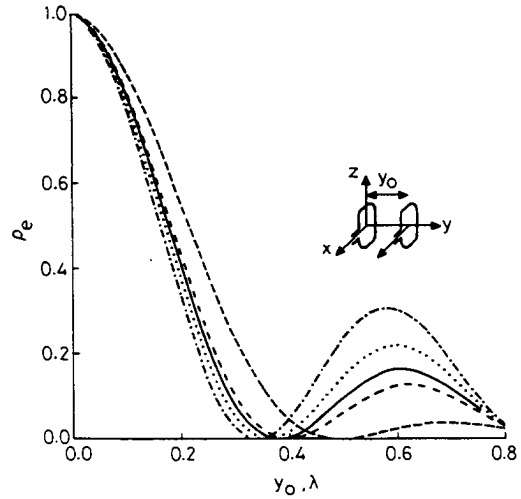


Fig. 11 *Envelope correlation coefficient against separation distance for loops of perimeter P with $b/a = 1$ and $v = 10$ compared with that of two antennas with isotropic patterns in horizontal plane*

——— isotropic
- - - - $P = 0.25\lambda$
------- $P = 0.50\lambda$
····· $P = 0.75\lambda$
—·—·— $P = 1.00\lambda$

rotation β of the second loop while the first loop is held stationary. As the antenna rotates, angle and polarisation

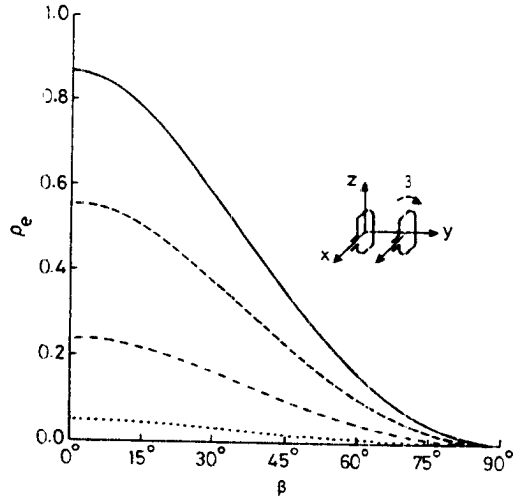


Fig. 12 *Envelope correlation coefficient against rotation angle for 0.25λ loops with b/a = 1 and v = 10 for various separation distances*

——— $y_0 = 0.1\lambda$
- - - - $y_0 = 0.2\lambda$
------- $y_0 = 0.3\lambda$
····· $y_0 = 0.4\lambda$

diversity result in a dramatic decrease in the signal correlation. Such an arrangement of antennas can be a useful means of achieving high diversity returns in mobile units where space is an important consideration, and in fact has been implemented in some cases [17, p. 148].

6 Conclusion

In this work we have presented the development of a computational model of two coupled loop radiators with arbitrary relative orientation and position. Use of the superquadric curve to represent the loop geometries allows investigation of a large range of practical antennas, including circular, elliptical, and rectangular loops, using one general analysis. A Galerkin method of moments technique with piecewise sinusoidal subsectional basis functions has been applied to solve the electric field integral equation for thin wires using a parametric representation of the superquadric loop. This construction allows the antenna to be represented with curved rather than piecewise linear subsectional segments, resulting in a computationally efficient algorithm. The current distribution in the loops computed from the integral equation is subsequently used to obtain the antenna self, mutual, and input impedance and directivity pattern. Several representative results have been given which not only provide insight into the effects of such factors as the loop squareness and aspect ratio on the antenna behaviour but also show the model's versatility in analysing many different types of structures. These examples further allow determination of the effect of placing the two loops in close proximity to each other as might occur in mobile communications applications. In many cases where the system configuration places constraints on the antenna shape, size, and position, incorporation of this formulation in the design process can help in determining the most suitable configuration to meet the needs of the system.

We have also discussed the concept of antenna diversity and have presented the expression for computing the envelope correlation coefficient for signals received by multiple antennas. This expression was applied to several coupled loop arrangements to illustrate the potential diversity performance of these radiators when used in a proper configuration. We have shown that with proper orientation and spacing, good diversity action can be obtained even for antennas which are spaced closely together. This information is crucial to the development of robust mobile receiver units designed for use in a multipath environment.

7 References

1 POCKLINGTON, H.C.: 'Electrical oscillations in wires', *Proc. Cambridge Phil. Soc.*, 1897, **9**, pp. 324–332
2 HALLÉN, E.: 'Theoretical investigations into transmitting and receiving qualities of antennae', *Nova Acta Regiae Soc. Sci. Ups., Ser. 4*, 1938, **11**, pp. 1–44
3 STORER, J.E.: 'Impedance of thin-wire loop antennas', *Trans. Am. Inst. Electr. Eng.*, 1956, **75**, pt. 1, pp. 606–619
4 WU, T.T.: 'Theory of the thin circular loop antenna', *J. Math. Phys.*, 1962, **3**, pp. 1301–1304
5 HARRINGTON, R.F., and MAUTZ, J.: 'Electromagnetic behaviour of circular wire loops with arbitrary excitation and loading', *Proc. IEE*, Jan. 1968, **115**, (1), pp. 68–77
6 ZHOU, G., and SMITH, G.S.: 'An accurate theoretical model for the thin-wire circular half-loop antenna', *IEEE Trans.*, 1991, **AP-39**, (8), pp. 1167–1177
7 ADACHI, S., and MUSHIAKE, Y.: 'Theoretical formulation for circular loop antennas by integral equation method', *Sci. Rep. Res. Inst. Tohoku Univ. Ser. B*, 1957, **9**, pp. 9–18
8 IIZUKA, K., KING, R.W.P., and HARRISON, C.W.: 'Self- and mutual admittances of two identical circular loop antennas in a conducting medium and in air', *IEEE Trans.*, 1966, **AP-14**, (4), pp. 440–450
9 ABUL-KASSEM, A.S., and CHANG, D.C.: 'On two parallel loop antennas', *IEEE Trans.*, 1980, **AP-28**, (4), pp. 491–496
10 TSUKIJI, T., and TOU, S.: 'On polygonal loop antennas', *IEEE Trans.*, 1980, **AP-28**, (4), pp. 571–575
11 BAKER, K., and LAGRONE, A.H.: 'Digital computation of the mutual impedance between thin dipoles', *IRE Trans.*, 1962, **AP-10**, pp. 172–178
12 BARR, A.H.: 'Superquadrics and angle-preserving transformations', *IEEE Comput. Graph. Appl.*, 1981, **1**, pp. 11–23
13 HARRINGTON, R.F.: 'Field computation by moment methods' (Macmillan, New York, 1968)
14 POGGIO, A.J., and MILLER, E.K.: 'Techniques for low-frequency problems', *in* LO, Y.T., and LEE, S.W. (Eds.): 'Antenna handbook: theory, applications, and design' (Van Nostrand Reinhold, New York, 1988), Chap. 3
15 CHAMPAGNE, N.J., WILLIAMS, J.T., and WILTON, D.R.: 'The use of curved segments for numerically modeling thin-wire antennas and scatterers', *IEEE Trans.*, 1992, **AP-40**, (6), pp. 682–689
16 VAUGHAN, R.G., and ANDERSEN, J.B.: 'Antenna diversity in mobile communications', *IEEE Trans. Veh. Technol.*, 1987, **VT-36**, (4), pp. 149–172
17 JAKES, W.C.: 'Microwave mobile communications' (Wiley, New York, 1974)
18 LEE, W.C.Y.: 'Mobile communications engineering' (Wiley, New York, 1982)
19 CLARKE, R.H.: 'A statistical theory of mobile radio reception', *Bell Syst. Tech J.*, 1969, **47**, pp. 957–1000
20 RAHMAT-SAMII, Y.: 'Useful coordinate transformations for antenna applications', *IEEE Trans.*, 1979, **AP-27**, pp. 571–574
21 TSAI, L.L.: 'A numerical solution for the near and far fields of an annular ring of magnetic current', *IEEE Trans.*, 1972, **AP-20**, (5), pp. 569–576

8 Appendix

8.1 Superquadric parametrisation

To permit the use of curved subsectional wire segments in the moment method analysis of superquadric loop antennas, it is necessary to first parametrise the curve in a convenient form. Among the many parametrisations possible, observation has shown that a form which provides mathematical simplicity and numerical stability uses the parameter τ and assumes the form

$$x = a\psi(\tau)\cos\tau \tag{8}$$

$$y = b\psi(\tau)\sin\tau \tag{9}$$

where

$$\psi(\tau) = \frac{1}{(|\sin\tau|^{\nu} + |\cos\tau|^{\nu})^{1/\nu}} \tag{10}$$

and $0 \leqslant \tau \leqslant 2\pi$. The unit tangent vector $\hat{s}$ in the direction of the perimeter co-ordinate s is

$$\hat{s} = \frac{-a|\sin\tau|^{\nu-1}\,\mathrm{sgn}\,(\sin\tau)}{\gamma(\tau)}\hat{x} + \frac{b|\cos\tau|^{\nu-1}\,\mathrm{sgn}\,(\cos\tau)}{\gamma(\tau)}\hat{y} \tag{11}$$

where

$$\gamma(\tau) = \sqrt{(a^2|\sin\tau|^{2\nu-2} + b^2|\cos\tau|^{2\nu-2})} \tag{12}$$

The differential arc length ds in terms of the differential $d\tau$ is expressed by

$$ds = |r_\tau|\,d\tau = \gamma(\tau)\psi^{\nu+1}(\tau)\,d\tau = \Delta(\tau)\,d\tau \tag{13}$$

Characterization of Antennas for Personal Wireless Communications Applications

M.A. Jensen and Y. Rahmat-Samii

International Journal of Wireless Information Networks
vol. 1, no. 3, pp. 165-176
March 1994

International Journal of Wireless Information Networks, Vol. 1, No. 3, 1994

Characterization of Antennas for Personal Wireless Communications Applications

Yahya Rahmat-Samii[1,2] and Michael A. Jensen[1]

The advancement of antenna technology in personal wireless communication systems has been encouraged by the increasingly stringent demands placed upon these systems to provide low-power and highly reliable information transfer. The antenna designer must not only consider the cost, manufacturability, compactness, and system integrability of the radiator but also generate a product which satisfies rigid specifications concerning return loss, bandwidth, and gain while operating in a complex radiating environment. Successful, cost-effective approaches to the design of antennas for communication devices rely upon the implementation of sophisticated analysis tools, such as the finite-difference time-domain (FDTD) method, capable of predicting the electromagnetic behavior of complicated topologies. In this work, the behavior of planar inverted F, monopole, and loop antennas is investigated using tools based upon the FDTD approach. Such factors as the effects of the conducting chassis, plastic casing, and biological tissue on the antenna performance are investigated. Experimental measurements are used to validate the results obtained from computations and to provide further insight into the behavior of the different geometries. The use of antenna diversity to reduce the effects of multipath fading is discussed, and several examples of antenna diversity configurations are provided.

KEY WORDS: Wireless; FDTD; PIFA; monopole; loop antenna; diversity.

1. INTRODUCTION

Personal wireless communication systems have been evolving steadily due to the increasingly stringent demands placed upon them to provide low-power, highly reliable information transfer with ergonomical designs. The antenna, being such a fundamental part of the overall unit, has certainly not been immune from this technological growth. New applications demand that the antenna designer not only consider the cost, manufacturability, compactness, and system integrability of the radiator but also generate a product which satisfies rigid specifications while operating in a complex radiating environment. Specifically, the antenna elements must ideally offer the features:

- Omnidirectionality to ensure reliable communication regardless of antenna orientation
- Wide bandwidth to accommodate spread-spectrum communication schemes
- Minimal coupling to the operator's biological tissue to avoid performance degradation and health risks
- Low-profile physical geometry to facilitate antenna packaging and provide convenience to the operator

This work presents the results of recent efforts aimed toward the analysis, design, and performance evaluation of antenna structures suitable for implementation in hand-held transceiver units for personal communications applications.

The intricate material and geometrical features as-

[1] Department of Electrical Engineering, University of California, Los Angeles, Los Angeles, California 90024-1594.
[2] Correspondence should be directed to Yahya Rahmat-Samii, Department of Electrical Engineering, University of California, Los Angeles, Los Angeles, California 90024-1594.

1068-9605/94/0700-0165$07.00/0

sociated with antenna configurations often cannot be practically modeled using conventional approximate analytical techniques. For this reason, cost-effective antenna design approaches must rely upon the implementation of sophisticated analysis tools capable of predicting the electromagnetic behavior of the antenna geometry while incorporating the effects of neighboring bodies such as a transceiver chassis or biological tissue. The finite-difference time-domain (FDTD) algorithm [1–2] provides a means of performing the detailed simulations required, and results of several efforts to model simple antenna systems have appeared in the literature [3–7]. In this work, the performance of several antenna designs suitable for hand-held transceiver units is investigated using computational tools based upon the FDTD approach [8]. The antenna elements considered include the monopole, the planar inverted F antenna, and the loop antenna. Such issues as the effect of the handset conducting chassis and plastic casing and the influence of operator proximity on the antenna behavior are addressed. Experimental verification of the simulated results for these antenna configurations is included to provide further evaluation of the performance of antennas designed for this application.

The concept of antenna diversity [8–12] in which multiple antennas are placed upon a single transceiver unit to increase the received signal-to-noise ratio is also discussed, and the feasibility of incorporating such designs into small portable transceivers and telephones is addressed. Several computational results are presented which evaluate the theoretically achievable improvements obtainable from antenna diversity configurations. Conclusions are drawn based upon the theoretical and experimental results.

2. PHYSICAL MODELING OF ANTENNA CONFIGURATIONS

2.1. Hand-Held Transceiver Geometry

Numerous different geometries and configurations of small antennas suitable for implementation in portable handsets are conceivable. For example, Fig. 1a depicts a typical handset with the monopole or whip antenna mounted on the top panel. This particular configuration has been used extensively in wireless communications applications—mainly due to its simplicity—and as such has been the focus of many theoretical and computational studies [3, 4]. Figure 1b illustrates the handset with a portion of the outer plastic casing removed to show the mounting of the dual planar inverted F antennas (PIFA) [13–15] placed on the handset sides. Figure 1c is a similar example for a loop antenna mounted on the back of the handset. These configurations have received relatively little attention in the

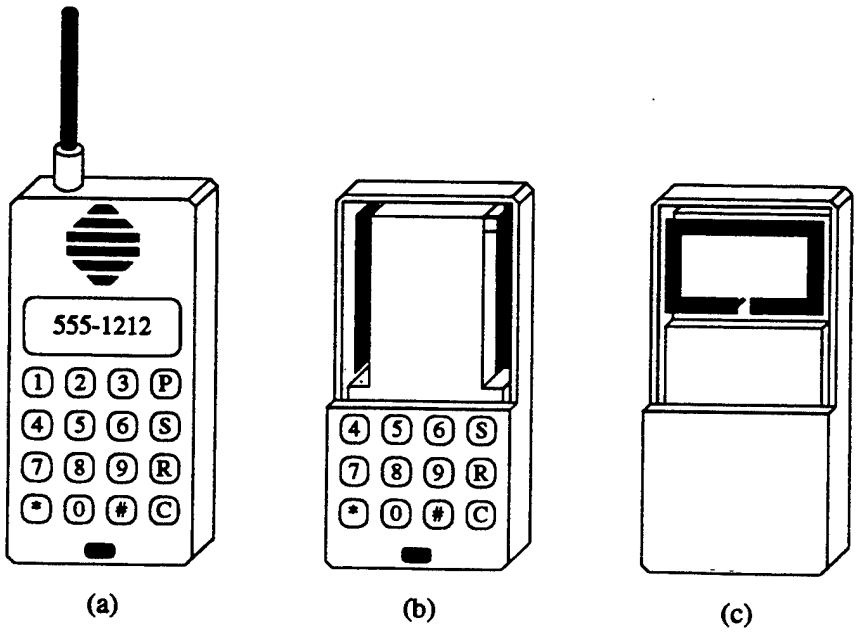


Fig. 1. Typical antenna configurations for hand-held communications devices: (a) monopole; (b) PIFA diversity antennas; (c) loop. The plastic casing has been partially removed from the front of (b) and the back of (c) to reveal the position of the antennas.

past, despite the fact that they may be conveniently packaged and protected by the plastic casing which encloses the device. However, new efforts to provide low-profile, unobtrusive antennas and to mount multiple antennas on a single transceiver unit in a diversity scheme have resulted in a recent effort to understand the performance of these structures.

In order to develop a high-performance design utilizing antenna geometries such as those depicted in Fig. 1, it is essential to evaluate the antenna behavior when placed in its operating environment. For example, the effect of the shielding chassis on which the antennas are mounted and the plastic casing surrounding the entire device must be investigated. Additionally, the effect of biological tissue near the antenna may have a pronounced effect on the antenna's electromagnetic characteristics. An understanding of these effects is gained through detailed computational studies and experimental investigations.

2.2. FDTD Implementation

The antennas depicted in Fig. 1 consist of wires, planar conducting surfaces, and dielectrics mounted in combination on a conductive chassis. Accurate electromagnetic modeling of such a configuration requires the use of sophisticated numerical techniques such as the FDTD method [1]. This approach is based upon a discretization of Maxwell's curl equations using centered difference approximations for the derivatives in both space and time, resulting in a set of algebraic equations for determining the field distributions in a given region. For readers not familiar with the implementation of this method, the key steps of the FDTD approach are summarized to demonstrate the simplicity and versatility of the method.

The starting point in the development of the FDTD technique is Maxwell's time-domain curl equations for an isotropic medium

$$\nabla \times \vec{E} = -\mu \frac{\partial \vec{H}}{\partial t} - \vec{M}$$

$$\nabla \times \vec{H} = \epsilon \frac{\partial \vec{E}}{\partial t} + \sigma \vec{E} + \vec{J} \qquad (1)$$

where ϵ, μ, and σ are the medium permittivity, permeability, and conductivity respectively and $\vec{E}$ and $\vec{H}$ are the electric and magnetic field intensities, respectively. The magnetic and electric current densities $\vec{M}$ and $\vec{J}$ represent source currents within the computational domain and will be set to zero in this implementation. Centered difference approximations to the space and time derivatives lead to the forms

$$H^{n+1/2}_{x,i,j,k} = H^{n-1/2}_{x,i,j,k} + \gamma \left[\frac{E^n_{y,i,j,k} - E^n_{y,i,j,k-1}}{\Delta z} - \frac{E^n_{z,i,j,k} - E^n_{z,i,j-1,k}}{\Delta y} \right]$$

$$H^{n+1/2}_{y,i,j,k} = H^{n-1/2}_{y,i,j,k} + \gamma \left[\frac{E^n_{z,i,j,k} - E^n_{z,i-1,j,k}}{\Delta x} - \frac{E^n_{x,i,j,k} - E^n_{x,i,j,k-1}}{\Delta z} \right]$$

$$H^{n+1/2}_{z,i,j,k} = H^{n-1/2}_{z,i,j,k} + \gamma \left[\frac{E^n_{x,i,j,k} - E^n_{x,i,j-1,k}}{\Delta y} - \frac{E^n_{y,i,j,k} - E^n_{y,i-1,j,k}}{\Delta x} \right]$$

$$E^{n+1}_{x,i,j,k} = \frac{1}{\beta} \left[\alpha E^n_{x,i,j,k} + \frac{H^{n+1/2}_{z,i,j+1,k} - H^{n+1/2}_{z,i,j,k}}{\Delta y} - \frac{H^{n+1/2}_{y,i,j,k+1} - H^{n+1/2}_{y,i,j,k}}{\Delta z} \right]$$

$$E^{n+1}_{y,i,j,k} = \frac{1}{\beta} \left[\alpha E^n_{y,i,j,k} + \frac{H^{n+1/2}_{x,i,j,k+1} - H^{n+1/2}_{x,i,j,k}}{\Delta z} - \frac{H^{n+1/2}_{z,i+1,j,k} - H^{n+1/2}_{z,i,j,k}}{\Delta x} \right]$$

$$E^{n+1}_{z,i,j,k} = \frac{1}{\beta} \left[\alpha E^n_{z,i,j,k} + \frac{H^{n+1/2}_{y,i+1,j,k} - H^{n+1/2}_{y,i,j,k}}{\Delta x} - + \frac{H^{n+1/2}_{x,i,j+1,k} - H^{n+1/2}_{x,i,j,k}}{\Delta y} \right] \qquad (2)$$

where

$$\alpha = \frac{\epsilon}{\Delta t} - \frac{\sigma}{2}, \qquad \beta = \frac{\epsilon}{\Delta t} + \frac{\sigma}{2}, \qquad \gamma = \frac{\Delta t}{\mu}. \qquad (3)$$

The superscripts in Eqs. (2) denote the instant of time ($t = n\Delta t$), and the subscripts denote the field component polarization (x, y, or z) as well as its spatial position ($x = i\Delta x$, $y = j\Delta y$, $z = k\Delta z$). This set of equations applies to a computational grid formed by filling a volume with the unit cells depicted in Fig. 2, an arrangement which ensures continuity of tangential fields across cell boundaries. Implementation of this algorithm provides a means for tracking the time evolution of the electromagnetic fields surrounding the antenna for a given set of initial and boundary conditions.

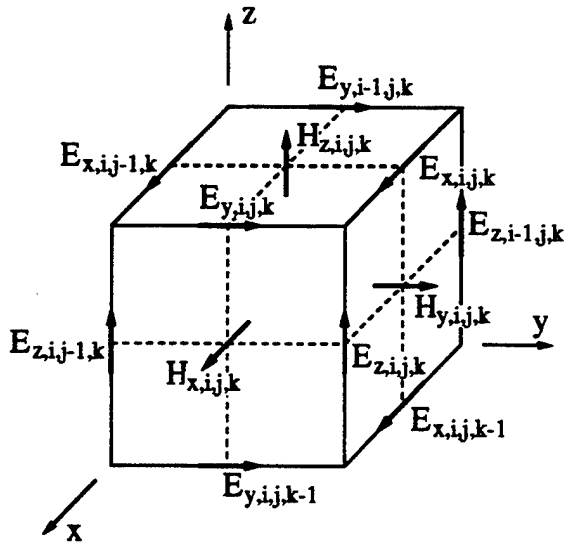


Fig. 2. Discretization of electromagnetic field components for the FDTD methodology.

To use the time-stepping equations in Eqs. (2), the general procedure outlined in Fig. 3 is followed. Initially, all field values within the computational domain are set to zero. The antenna excitation is then introduced by specifying the antenna voltage (electric field) at the antenna feed point. The value of $\vec{H}$ is then computed at time $t = (n + 1/2)\Delta t$, followed by the computation of $\vec{E}$ at time $t = (n + 1)\,\Delta t$. Boundary conditions on the perfect electric conductors are enforced by setting the tangential components of $\vec{E}$ to zero at mesh points coincident with the conductors. Special time-stepping equations are used in the vicinity of thin wires to account for the finite wire radius. The time-stepping procedure is repeated until the desired response for the antenna has been obtained. Frequency-domain quantities are then obtained from the time-domain data through use of a Fourier transform generally in the form of a fast Fourier transform (FFT).

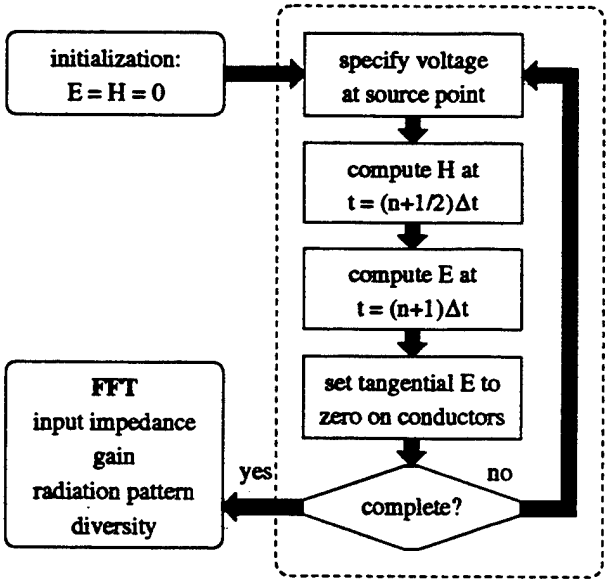


Fig. 3. Practical implementation of the FDTD algorithm to track the time evolution of electromagnetic fields surrounding the antenna.

Theoretically, the simulation grid surrounding the antenna under investigation must extend to infinity for perfect representation of the actual radiation scenario. However, computational limitations require that the grid be truncated some finite distance from the radiator, resulting in artificial reflections which can substantially degrade the accuracy of the computed results. Approximate absorbing boundary conditions (ABCs) [16] are therefore implemented on the outer boundary of the computational grid such that fields incident on the truncation planes will be absorbed rather than reflected. Since these ABCs are approximations, it is necessary to include an appropriate space between the antenna under investigation and the computational grid truncation planes. Through observation, we have found this distance to range from $\lambda/4$ to $\lambda/2$.

The FDTD algorithm as posed is an explicit difference method whose stability is governed by the Courant–Friedrichs–Lewy condition [1], namely

$$\left(\frac{1}{\Delta x^2} + \frac{1}{\Delta y^2} + \frac{1}{\Delta z^2}\right)^{1/2} \Delta t \leq \frac{1}{v_{\max}} \qquad (4)$$

where $v_{\max}$ is the velocity of light in the medium with the smallest permittivity within the computational domain. For most cases, this will be free space unless the entire domain is filled with dielectric.

3. COMPUTATIONAL AND EXPERIMENTAL RESULTS

The following cases provide examples of computations and measurements used in the design and evaluation of different antenna structures for wireless handheld transceivers. The experimental measurements of S_{11} provided here have been obtained from a Hewlett-Packard 8510B network analyzer at the University of California, Los Angeles antenna measurement facility. This parameter is a measure of the reflection coefficient for a voltage wave introduced to the antenna feed point using a 50-Ω coaxial cable. In many examples, only the magnitude of S_{11} is shown since such a plot readily displays

the degree to which the antenna is matched to a feed line and the antenna bandwidth (a perfect match is $|S_{11}| \rightarrow -\infty$ dB). For each of the antennas under investigation, a wire is used either simply as a feeding mechanism or as the antenna itself in the case of the monopole. In each case, the wire used has a radius r_0 given by

$$2r_0 = 0.9195 \text{ mm} \tag{5}$$

which corresponds to the inner conductor diameter for RG402/U semirigid 50-Ω coax.

The geometries and dimensions of the monopole, planar inverted F, and loop antennas on the handset are depicted in Fig. 4. The conducting box dimensions provided in Fig. 4a are valid for all three configurations.

3.1. Monopole

As mentioned, the monopole, or whip antenna as it is commonly called, has become a widely used radiator for wireless communication networks because of its simplicity and desirable electrical properties. This fact has naturally led to numerous analytical models of the whip antenna placed upon an infinite or a finite conducting plate. However, only very recently has research been focused on the effect of the transceiver handset or the operator's hand near the antenna on its electromagnetic characteristics.

The radiation patterns normalized to the antenna gain for the monopole on the handset shown in Fig. 4a are provided in Fig. 5 at a frequency of 915 MHz where the monopole length is approximately a quarter wavelength. These patterns are compared with those for a half-wavelength dipole antenna, which illustrates the familiar "doughnut" pattern. As can be seen, the pattern for the monopole on the conductive chassis differs considerably from the simple dipole pattern. Most notably, the direction of peak gain is shifted to an angle below the horizontal plane. Furthermore, the asymmetric placement of the monopole on the handset produces an asymmetrically shaped pattern and introduces a cross-polarized component into the radiation pattern.

Figure 6 compares the magnitude of the reflection parameter S_{11} versus frequency for the whip/handset configuration with and without a 6.56-mm lossless dielectric casing with $\epsilon_r = 2$ which completely surrounds the chassis. It should be noted that the thickness of this casing has been chosen to coincide with the cell size for the FDTD computation, and is actually somewhat thicker than practical casings found on modern handsets. This result shows that the effect of the plastic casing on the antenna behavior becomes more pronounced as the frequency increases since the dielectric thickness becomes electrically larger at these frequencies. Also included in the figure is the measured value of S_{11} for an experimental version of the design constructed in our laboratory without the plastic casing. The agreement between the computed and measured results is seen to be good. The disagreement at low frequencies is most likely

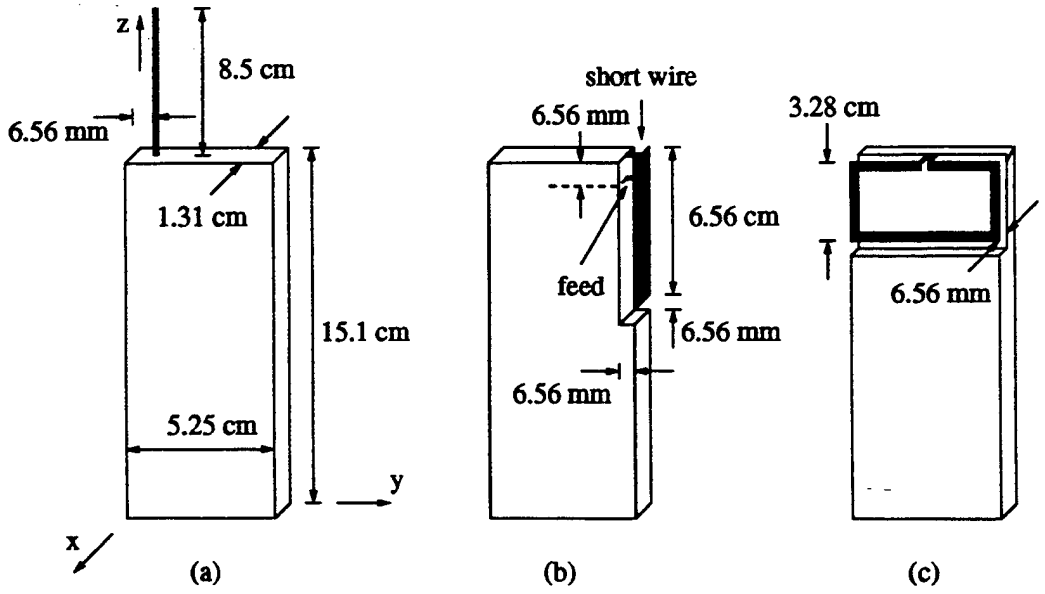


Fig. 4. Geometries and dimensions for the antennas on the hand-held transceiver: (a) monopole; (b) side-mounted PIFA; (c) back-mounted strip loop antenna of strip width 6.56 mm.

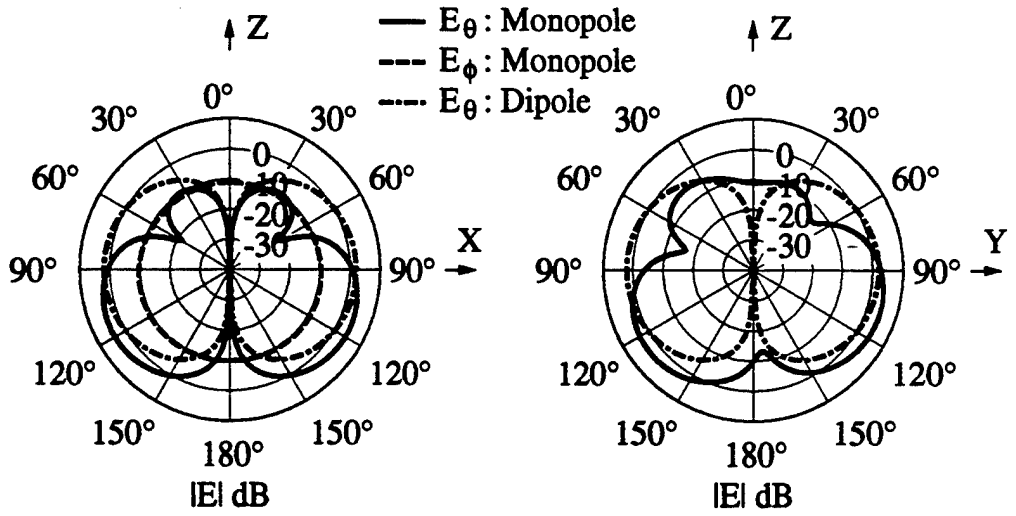


Fig. 5. Gain patterns at 915 MHz for the monopole on the handset compared to that of a half-wavelength dipole.

due to the fact that pronounced coupling at longer wavelengths between the antenna and the absorbing boundaries in the computations and nearby objects in the experiments result in reflections which can corrupt the impedance data.

As another example of the types of computations and measurements used in the design of the monopole antenna, attention is turned to the effect of biological tissue on the antenna performance. The antenna configuration shown in Fig. 4a is augmented with a lossy dielectric sheet of thickness 1.97 cm covering three sides of the conducting chassis to represent a human hand holding the antenna, as indicated in the figure inset in Fig. 7. The dielectric is characterized by $\epsilon_r = 51$ and $\sigma = 1.6$ S/m, values which represent a typical human muscle at frequencies near 915 MHz [17]. The S_{11} data for this computation, as well as that obtained experimentally, is shown in Fig. 7. As can be seen, the simulated results agree quite well with experiment, and both results show that the presence of the human hand serves to lower the resonant frequencies of the antenna. Experimental data are also provided for the case of the handset held at an angle of approximately 50° from its upright orientation near the human head. Again, the antenna resonant frequency is reduced by the presence of additional biological tissue. Because of space limitations in this paper, no computational examples of the antenna performance near a human head are included. However, a detailed study of the interaction between the head and the antenna will be provided in a future paper.

The gain patterns for the monopole antenna configuration with both the hand and the plastic casing men-

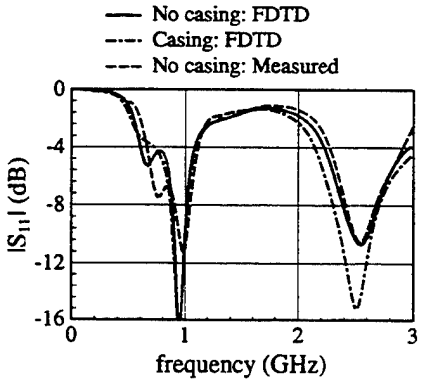


Fig. 6. Magnitude of S_{11} for the monopole on the handset with and without a plastic casing surrounding the conducting chassis.

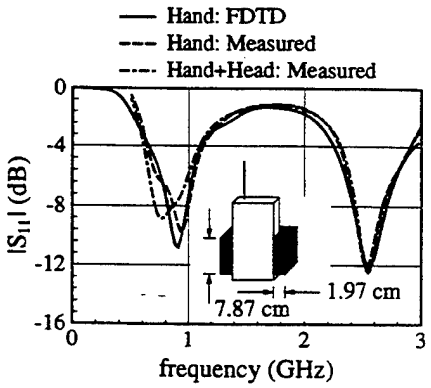


Fig. 7. Magnitude of S_{11} for the monopole on the handset with a human hand and head in proximity to the antenna/chassis structure.

tioned appear in Fig. 8. Comparison of these plots with those in Fig. 5 shows that the presence of these dielectrics alters both the shape and polarization characteristics of the antenna radiation patterns. Also noteworthy is the slight reduction in gain which results from absorption within the body. To quantify this effect, the power absorbed by the body is computed using the integration

$$P_{abs} = \tfrac{1}{2} \int_V \sigma |\mathbf{E}|^2 \, dV \tag{6}$$

over the volume of the hand, where **E** is the frequency domain value of the electric field. Using this result in conjunction with the value for the real power contained in the far-field pattern reveals that the hand absorbs approximately 20% of the total delivered power for this scenario. Computations reveal that over 50% of the delivered power is absorbed if the head is also near the handset.

3.2. Planar Inverted F Antenna

In contrast to the case of the monopole, elements such as the planar inverted F antenna (PIFA) have received relatively little attention. This interesting antenna topology is essentially an air-dielectric modified microstrip which allows a simple impedance match in a low-profile design [13–15]. The geometry is illustrated in Fig. 4b where the antenna is mounted on the conducting chassis. As can be seen, the PIFA has the advantage of being low profile and can therefore be conveniently packaged, as implied in Fig. 1b. The short-circuit introduced by the pin at one edge of the radiating element acts to approximately halve the structure size for a given resonant frequency. Additionally, by properly choosing the antenna dimensions, the element may be matched to a 50-Ω feeding coax. Although this element normally exhibits a very narrow bandwidth when placed on a large ground plane, it has been observed that its bandwidth increases considerably, sometimes becoming as high as 10%, when it is located on a small ground plane or conducting box [13].

The performance of the PIFA designed for a handheld transceiver may be evaluated using the FDTD methodology. For example, Fig. 9 shows a prediction of the antenna gain pattern for the structure shown in Fig. 4b at the design frequency of 915 MHz. As may be suspected due to the long, slender appearance of this particular PIFA configuration, the patterns resemble those of a monopole antenna and provide approximate omnidirectionality in the horizontal (xy) plane. However, the input impedance, which is plotted versus frequency in Fig. 10 over a band near the design frequency of 915 MHz, differs significantly from that of the monopole. As can be determined from this plot, the PIFA is very nearly matched to a 50-Ω feeding coaxial cable and can therefore be implemented without including a matching network. Figure 11 compares the magnitude of S_{11} for this design to measured data obtained in our laboratory. These results not only illustrate the capability of the simulation tools developed in predicting the antenna characteristics but also the good impedance match (low reflection coefficient) and reasonable bandwidth ($\sim 7.7\%$ for VSWR < 2) obtainable with this design.

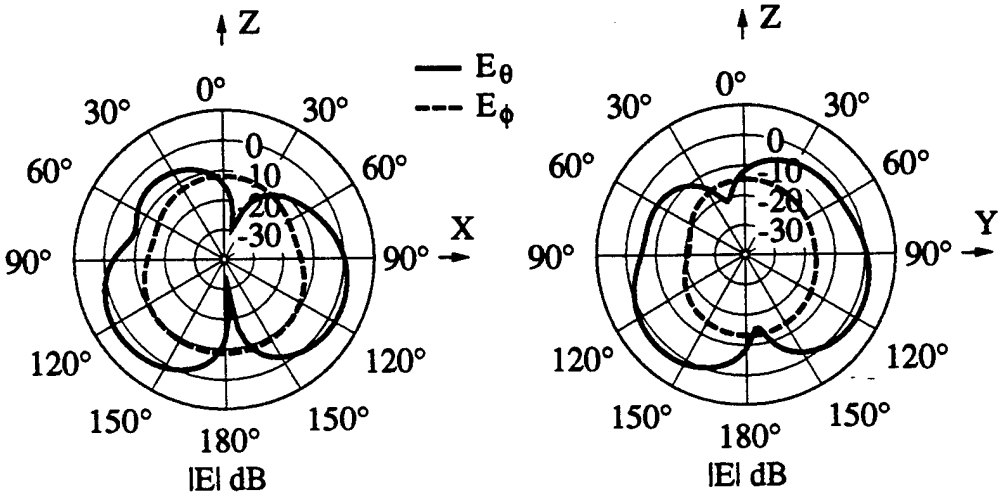


Fig. 8. Gain patterns at 915 MHz for the monopole on the handset with the hand and the plastic casing present.

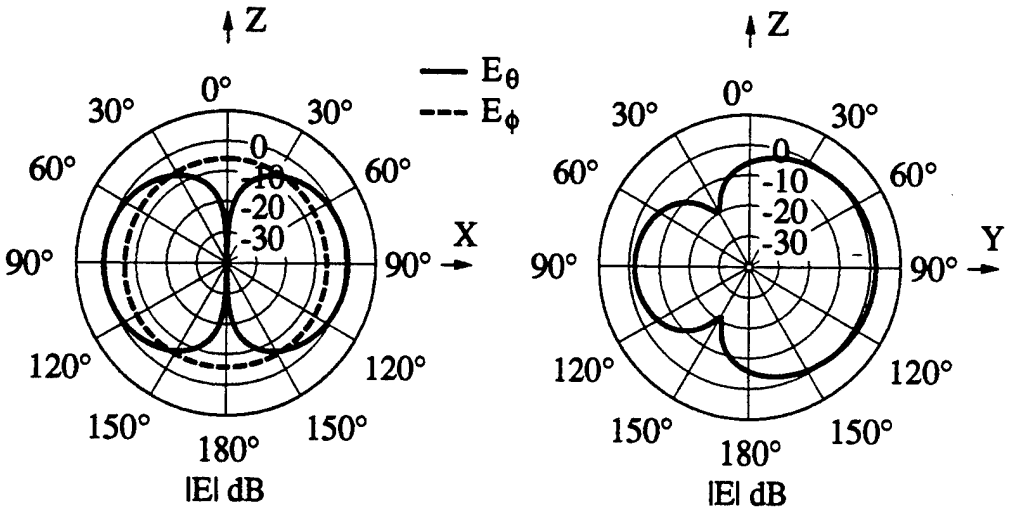


Fig. 9. Gain patterns at 915 MHz for the PIFA mounted on the transceiver.

Computations and experiments performed in our laboratory show that, aside from a slight decrease in the antenna resonance frequency, the input impedance of the design depicted in Fig. 4b is minimally affected by the addition of a human hand or head provided that the hand is not placed directly over any part of the antenna. Therefore, hardware which incorporates such a design must provide an outer plastic casing which encourages the user to place his hand below the antenna if reliable communication is to be obtained. This plastic casing will lower the resonance frequency of the antenna, which will in turn require that the antenna structure be tuned such that the final design resonates at the desired operating frequency. Additionally, it is expected that the radiation pattern will be noticeably altered by the presence of the hand, head, and plastic casing. Results from detailed analyses performed in our laboratory as well as conclusions drawn from our findings will appear in a subsequent communication.

3.3. Loop Antenna

In some cases, emerging technologies in RF circuitry design are dictating that balanced push-pull amplifiers be used to drive the antenna. For these circuit configurations, the use of balanced antennas such as loops or printed dipoles allows the interface between the antenna and amplifier without requiring a balanced-to-unbalanced transition (balun). A low-profile balanced antenna design appears in Fig. 4c where a strip loop antenna on the transceiver handset is illustrated [9]. In this case, the strip width is 6.56 mm and the loop is placed on the side of the handset which faces away from the operator. The gain patterns at 915 MHz for this configuration provided in Fig. 12 show that, unlike the PIFA and monopole antennas considered above, the predominant field polarization of this geometry changes between the two principal planes. Additionally, the approximate omnidirectionality of the patterns in Fig. 12

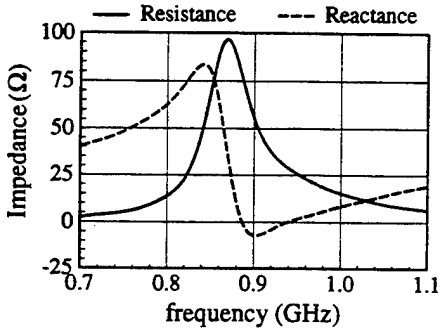


Fig. 10. Input impedance vs. frequency for the PIFA mounted on the transceiver.

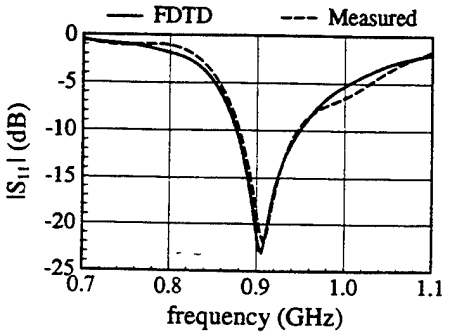


Fig. 11. Comparison of the computationally and experimentally obtained magnitude of S_{11} vs. frequency for the PIFA mounted on the transceiver.

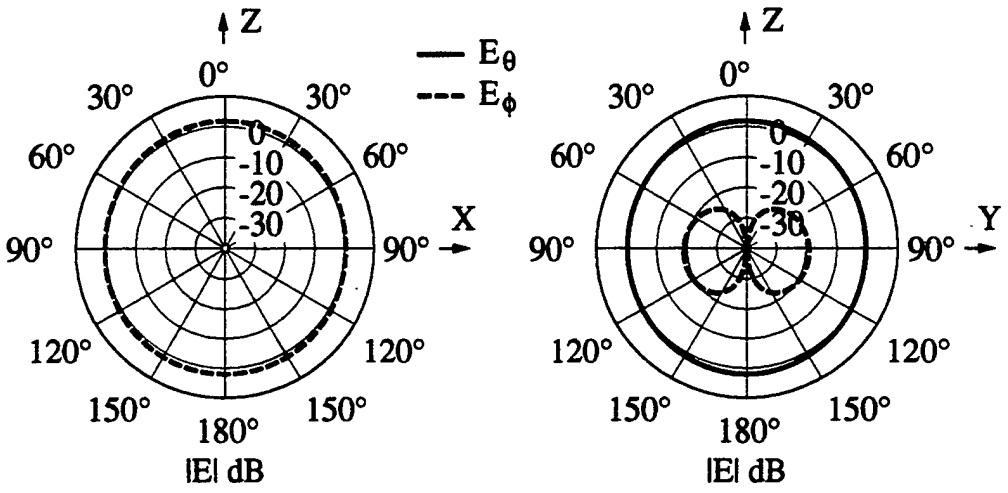


Fig. 12. Gain patterns at 915 MHz for the strip loop mounted on the handset.

shows that the performance of this antenna should be changed very little as the device orientation is changed during operation. The impedance versus frequency plot for this configuration given in Fig. 13 implies that matching the antenna to the feed line presents a considerable challenge to the designer. Note that this is in contrast to the well-matched nature of the PIFA antenna impedance shown in Fig. 10. However, significant efforts have recently been devoted to the development of novel feeding and loading techniques for the loop [9, 18] which properly adjust the impedance presented at the antenna terminals. Once this issue is satisfactorily addressed, the loop antenna becomes a simple and useful radiator for personal wireless devices.

4. ANTENNA DIVERSITY

4.1. Formulation

As demands increase for modern communication systems which provide reliable, robust information transfer even in fading environments, sophisticated diversity techniques must be implemented to combat the effects of multipath propagation. This diversity may be accomplished either by using additional frequency bands in spread-spectrum techniques or by using multiple "diverse" antennas in a single frequency band. To date, the additional receiver hardware complexity and increased power consumption required to implement antenna diversity has discouraged its use in portable equipment. However, new efforts to improve communication reliability *without* consuming additional spectrum is leading to the increased use of antenna diversity in mobile systems.

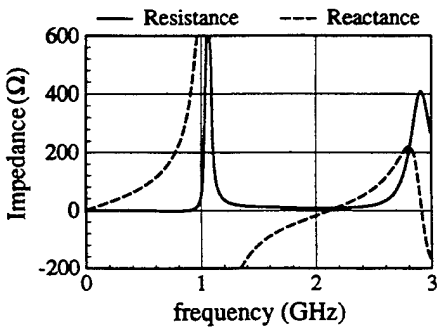


Fig. 13. Input impedance vs. frequency for the strip loop mounted on the handset.

Evaluation of the performance of multiple antennas configured in a diversity arrangement requires investigation of the statistics of the signals received by the elements through the envelope correlation coefficient ρ_e. An analytical expression for this quantity, which provides a measure of the similarity of the voltages at each antenna terminal, can be obtained if proper assumptions concerning the nature of the incoming multipath signal are made. These assumptions are as follows:

- The Rayleigh probability density function describes the envelope of the fading signal.
- Orthogonal polarizations in the incoming multipath wave are uncorrelated.
- Each individual polarization is spatially uncorrelated (i.e., wave incoming at (θ, ϕ) is uncorrelated with wave at (θ', ϕ') for $\theta' \neq \theta$ or $\phi' \neq \phi$).
- The spatial distribution of the incoming multipath waves is limited to the horizontal plane only ($\theta = \pi/2$) and the mean power density for all incident angles within this plane is constant.

Using these assumptions leads to an expression for the envelope correlation coefficient for the signals received by the two antennas which assumes the form [11]

$$\rho_e = \frac{\left| \int_0^{2\pi} A_{12}(\phi)\, d\phi \right|^2}{\left| \int_0^{2\pi} A_{11}(\phi)\, d\phi \int_0^{2\pi} A_{22}(\phi)\, d\phi \right|} \tag{7}$$

where

$$A_{pq}(\phi) = E_{\theta p}(\pi/2, \phi) E^*_{\theta q}(\pi/2, \phi) + \Gamma^2 E_{\phi p}(\pi/2, \phi) E^*_{\phi q}(\pi/2, \phi) \tag{8}$$

and $\vec{E}_p(\theta, \phi) = E_{\theta p}(\theta, \phi)\hat{\theta} + E_{\phi p}(\theta, \phi)\hat{\phi}$ is the vector radiation pattern associated with antenna p. The parameter Γ is the mean discrimination of the cross-polarized components [10, 19] given by

$$\Gamma = \frac{|\overline{E}^{inc}_{\phi}|}{|\overline{E}^{inc}_{\theta}|} \tag{9}$$

where $\overline{E}^{inc}_{\phi}$ and $\overline{E}^{inc}_{\theta}$ are the mean field strengths of the $\hat{\phi}$ and $\hat{\theta}$ components in the incident signal. The coordinate system assumed here has the z-axis aligned perpendicular to the plane of the earth.

Figure 14 shows the cumulative probability distribution of the signal-to-noise ratio (SNR) normalized to its average (⟨SNR⟩) for various values of ρ_e and for two-branch antenna diversity. Maximal ratio combining

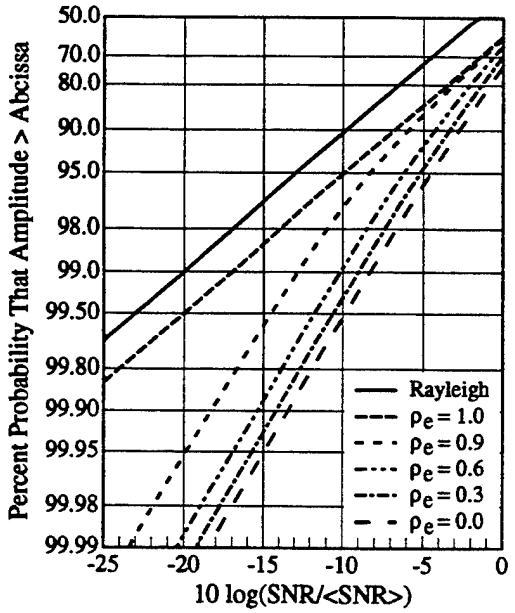


Fig. 14. Cumulative probability distribution for two-branch antenna diversity for different values of the envelope correlation coefficient ρ_e.

in which the two signals are received, cophased, properly weighted, and added is assumed in this plot. Also shown is the Rayleigh distribution which corresponds to a single antenna in the multipath environment. As can be seen, reduction of the envelope correlation coefficient provides a considerable increase in the probability of receiving a signal of adequate strength for reliable communication. For example, when no diversity is used, the Rayleigh distribution predicts that the received SNR will be above −20 dB 99% of the time. If two branch antenna diversity is used with $\rho_e = 0.6$, the SNR at the 99% reliability level increases to approximately −10 dB. The goal in diversity antenna design is therefore to minimize the envelope correlation coefficient to the extent possible using a combination of spatial, angle, and polarization diversity. A rule-of-thumb value for the maximum acceptable value of ρ_e for a diversity antenna configuration is around $\rho_e = 0.7$.

The method for computing the vector patterns $\vec{E}_1$ and $\vec{E}_2$ depends upon the type of diversity combining used on the received signals. For example, if switched or selection diversity is to be implemented, the pattern $\vec{E}_1$ should be determined as the pattern for element 1 radiating in the presence of element 2 which is open circuited. For other combining techniques, the pattern should be computed for element 2 terminated with a matched load (absorbing boundary in the FDTD computation). It is important to note that in these cases, each pattern is obtained for the respective element radiating in the presence of a *parasitic* rather than *excited* second antenna. In this manner, the simulation closely models the case of one antenna receiving a strong signal, while the second is positioned in a fade resulting from the multipath effect.

4.2. Examples

Determination of the diversity performance of two communications antennas mounted on a conducting handset requires computation of the envelope correlation coefficient provided in Eq. (7). To perform these computations, the antenna/handset coordinate system is assumed to be oriented such that the z-axis is 45° in the yz plane from the normal to the earth's surface. The incoming multipath wave is in the plane of the earth. As a first example of this computation, the PIFA geometry of Fig. 4b is used with a second, mirror image of the element placed on the opposite end of the box, as shown in the inset of Fig. 15. Each antenna is excited in turn while the other is terminated with an absorbing boundary at the source point. The FDTD technique is used to obtain the patterns for each element which are subsequently placed into Eq. (7) to obtain ρ_e. The resulting

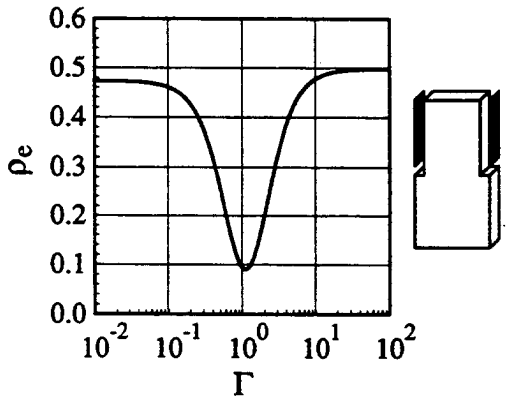


Fig. 15. Envelope correlation coefficient vs. mean polarization discrimination Γ for two PIFA antennas mounted on a transceiver handset. The handset is rotated 45° from upright in the *yz* plane.

values of ρ_e are plotted versus the mean polarization discrimination Γ from Eq. (9) in Fig. 15. Following the rule mentioned in the previous section that acceptable diversity is obtained when $\rho_e \leq 0.7$, these results show that acceptably low values of the correlation coefficient occur for all values of Γ. The value of ρ_e is minimized when there is a mixture of the two orthogonal polarizations, indicating that the polarization characteristics of the *total* field patterns of the two antennas offer considerable diversity performance. In most urban multipath scenarios, it is assumed that either polarization is equally likely (Γ = 1). A second interesting example is the combination of the monopole in Fig. 4a with the loop in Fig. 4c, shown in the inset of Fig. 16. Here again, the value of ρ_e is plotted versus Γ. As can be ascertained from this plot, the diversity performance is highest when both polarizations are equally likely in the incident field.

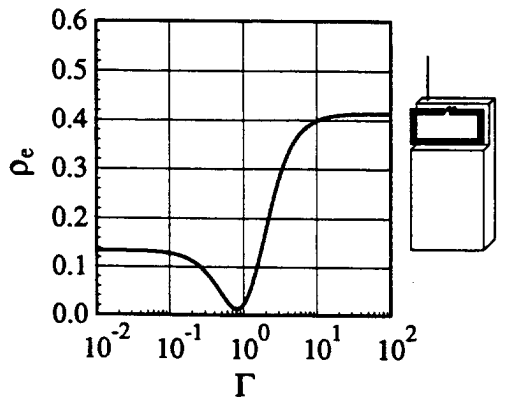


Fig. 16. Envelope correlation coefficient vs. Γ for a monopole and a loop antenna mounted on a transceiver handset. The handset is rotated 45° from upright in the *yz* plane.

5. CONCLUSIONS

This paper has presented some of the key procedures involved in the design of antenna structures for personal wireless communications devices. Concepts relating to the FDTD algorithm for electromagnetic analysis have been summarized, and advanced simulation tools based upon this methodology have been described. The electromagnetic behavior of several representative antenna configurations involving monopole, PIFA, and loop radiating elements has been predicted using the FDTD simulation tools. Additionally, the effects of the transceiver conducting chassis, the outer plastic casing, and nearby biological tissue such as the human hand and head on the performance of several topologies have been demonstrated. In many examples, experimental data has been presented which verifies the accuracy of the simulations and provides additional insight into the behavior of the different antennas.

The concept of antenna diversity has been addressed and some of the key assumptions and expressions relating to the statistics of the fields in a multipath fading environment have been reviewed. Two examples of feasible antenna diversity schemes for hand-held transceiver applications were included to illustrate the effectiveness of simple dual-antenna configurations in providing protection from signal fades due to multipath interference in urban and indoor communication scenarios.

ACKNOWLEDGMENTS

This work is funded by ARPA Contract DAAB07-93-C-C501.

REFERENCES

1. K. S. Yee, Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media, *IEEE Transactions on Antennas and Propagation*, Vol. AP-14, pp. 302–307, 1966.
2. A. Taflove and M. E. Brodwin, Numerical solution of steady-state electromagnetic scattering problems using the time-dependent Maxwell's equations, *IEEE Transaction on Microwave Theory and Techniques*, Vol. MTT-23, pp. 623–630, 1975.
3. R. Luebbers, L. Chen, T. Uno, and S. Adachi, FDTD calculation of radiation patterns, impedance, and gain for a monopole antenna on a conducting box, *IEEE Transactions on Antennas and Propagation*, Vol. 40, pp. 1577–1583, 1992.
4. J. Toftgard, S. N. Hornsleth, and J. B. Andersen, Effects on portable antennas of the presence of a person, *IEEE Transactions on Antennas and Propagation*, Vol. 41, pp. 739–746, 1993.
5. R. J. Luebbers and J. Beggs, FDTD calculation of wide-band antenna gain and efficiency, *IEEE Transactions on Antennas and Propagation*, Vol. 40, pp. 1403–1407, 1992.

6. J. G. Maloney, G. S. Smith, and W. R. Scott, Jr., Accurate computation of the radiation from simple antennas using the finite-difference time-domain method, *IEEE Transactions on Antennas and Propagation*, Vol. 38, pp. 1059–1068, 1990.
7. P. A. Tirkas and C. A. Balanis, Finite-difference time-domain method for antenna radiation, *IEEE Transactions on Antennas and Propagation*, Vol. 40, pp. 334–340, 1992.
8. M. A. Jensen and Y. Rahmat-Samii, FDTD analysis of PIFA diversity antennas on a hand-held transceiver unit, *1993 IEEE AP-S International Symposium Digest*, Ann Arbor, MI, vol. 2, pp. 814–817, 1993.
9. M. A. Jensen and Y. Rahmat-Samii, Isolated and coupled superquadric loop antennas for mobile communications applications, *Proceedings of the 3rd International Mobile Satellite Conference IMSC'93*, Pasadena, CA, pp. 581–586, 1993.
10. M. A. Jensen and Y. Rahmat-Samii, Characterization of electromagnetically coupled superquadric loop antennas for mobile communications applications, *IEE Proceedings H*. Vol. 141, pp. 85–93, 1994.
11. R. G. Vaughan and J. B. Andersen, Antenna diversity in mobile communications, *IEEE Transactions on Vehicular Technology*, Vol. VT-36, pp. 149–172, 1987.
12. W. C. Jakes, Jr., *Microwave Mobile Communications*, Wiley, New York, 1974.
13. K. Hirasawa and M. Haneishi, eds., *Analysis, Design, and Measurement of Small and Low Profile Antennas*, Artech House, Boston, 1992.
14. T. Taga and K. Tsunekawa, Performance analysis of a built-in planar inverted F antenna for 800 MHz band portable radio units, *IEEE Journal on Selected Areas in Communications*, Vol. SAC-35, pp. 921–929, 1987.
15. K. Kagoshima, K. Tsunekawa, and A. Ando, Analysis of a planar inverted F antenna fed by electromagnetic coupling, *1992 IEEE AP-S International Symposium Digest*, Chicago, pp. 1702–1705, 1992.
16. G. Mur, Absorbing boundary conditions for the finite-difference approximation of the time-domain electromagnetic field equations, *IEEE Transactions on Electromagnetic Compatibility*, Vol. EMC-23, pp. 377–382, 1981.
17. D. M. Sullivan, Frequency-dependent FDTD methods using Z-transforms, *IEEE Transactions on Antennas and Propagation*, Vol. 40, pp. 1223–1230, 1992.
18. K. Cho and K. Kagoshima, Monopole loaded small loop antennas eliminating a matching circuit, *1993 IEEE AP-S International Symposium Digest*, Ann Arbor, MI, vol. 3, pp. 1862–1865, 1993.
19. Y. Yamada, Y. Ebine, and N. Nakajima, Base station/vehicular antenna design techniques employed in high-capacity land mobile communications system, *Review of the ECL*, vol. 35, pp. 115–121, 1987.

Yahya Rahmat-Samii has been a professor of electrical engineering at the University of California, Los Angeles (UCLA) since 1988. He has held the position of senior research scientist at NASA's Jet Propulsion Laboratory, California Institute of Technology since 1978, where he contributed significantly to advanced antenna technology for space programs. He was a guest professor at the Technical University of Denmark (TUD) in the summer of 1986. He has also been a consultant to many aerospace companies. Dr. Rahmat-Samii received the M.S. and Ph.D. degrees in electrical engineering from the University of Illinois, Champaign-Urbana. Dr. Rahmat-Samii is a Fellow of IEEE (1985), a Fellow of IAE (1986), and the 1984 recipient of the Henry Booker Award of URSI. He was appointed an IEEE Antennas and Propagation Society Distinguished Lecturer and presented lectures internationally. He was an elected IEEE AP-S AdCom member for the second term and has been an associate editor of the *IEEE Antennas and Propagation Transactions* and magazine. In 1993, Dr. Rahmat-Samii was elected as the vice president of IEEE Antennas and Propagation Society. He was the Chairman of the IEEE Antennas and Propagation Society of Los Angeles from 1987 to 1989. In 1989, his chapter received the best Antennas and Propagation chapter award from the AP Society. He is one of the three international editors of the IEE book series on electromagnetics and antennas. He is also one of the editors of the *Journal of Electromagnetic Waves and Applications*. Dr. Rahmat-Samii has been one of the directors of AMTA (Antenna Measurement Techniques Association) since 1991 and is currently serving as its vice president. Yahya is listed in *Who's Who in America, Who's Who in Frontiers of Science and Technology*, and *Who's Who in Engineering*. Dr. Rahmat-Samii has authored or coauthored, over 300 technical journal articles and conference papers and has written chapters in 13 books. He has had pioneering contributions to diverse areas of antennas and electromagnetics. For these contributions he has received numerous NASA certificates of recognition and recently earned the JPL Team NASA's Distinguished Group Achievement Award. In 1992, he was the recipient of the Best Application Paper Award (Wheeler Award) for his paper published in the 1991 IEEE *Antennas and Propagation Transactions*. Dr. Rahmat-Samii is a member of Commissions A, B, and J of USNC/URSI, Sigma Xi, Eta Kappa Nu, and Electromagnetics Academy.

Michael A. Jensen was born in Fullerton, California, on March 27, 1966. He received the B.S. (summa cum laude) and M.S. degrees in electrical engineering from Brigham Young University in 1990 and 1991, respectively. He is currently pursuing the Ph.D. in electrical engineering at the University of California, Los Angeles. From 1989 to 1991 he was a graduate research assistant in the Lasers and Optics Laboratory at BYU. In 1990 he received a National Science Foundation Graduate Fellowship. Since 1991, he has been with the Antenna Laboratory at UCLA as a graduate student researcher. His main research interests include optical fiber communication, radiation and propagation for personal communications, and numerical electromagnetics. He is a member of Eta Kappa Nu and Tau Beta Pi.

A Low-Power CMOS Digitally Synthesized 0-13 MHz Sinewave Generator

G. Chang, A. Rofougaran, M.-K. Ku, A.A. Abidi, and H. Samueli

IEEE 1994 International Solid-State Circuits Conference Digest of Technical Papers
pp. 32-33
February 1994

WP 2.2: A Low-Power CMOS Digitally Synthesized 0-13MHz Agile Sinewave Generator*

Glenn Chang, Ahmadreza Rofougaran, Mong-Kai Ku, Asad A. Abidi, Henry Samueli

Electrical Engineering Department, Universuty of California, Los Angeles, CA

A low-power monolithic 1μm CMOS IC generates a 26MHz-wide single-sideband, frequency-hopped spread-spectrum waveform for wireless transmission in the 902-928MHz unlicensed ISM band. A direct digital frequency synthesizer (DDFS) on this IC produces 10b samples of sine and cosine waveforms whose frequency is selected with an 11b input word, followed by an on-chip 10b D/A converter (DAC) to convert the DDFS output into a sampled-data analog signal. Simplifications in architecture and circuits reduce dissipation of this 2.9x4.9mm² IC to 40mW at 40MHz with 3V supply.

A signed I-Q architecture for the frequency synthesizer halves the maximum clock frequency required to meet the specifications. Two DDFS/DAC channels produce discrete-time sine and cosine waveforms whose frequency varies from 0 to 13MHz (Figure 1). After anti-alias filtering, these sinewaves are respectively upconverted by quadrature outputs from a 915MHz local oscillator. If the two upconverted outputs are added, the output frequency ranges from 915 to 928 MHz. If they are subtracted, the frequency will cover 902 to 915MHz. Sign reversal may be implemented in the digital domain on the DDFS output. The DDFS clock must be at least three times the highest sinewave frequency to reduce anti-alias filter complexity. This DDFS/DAC thus must clock at least at 39 MHz.

The DDFS cell uses an efficient architecture to synthesize sinewaves with low spectral impurity [1]. The smallest internal and external word lengths that guarantee spurious tones at least -72.6dBc relative to the fundamental frequency are used. In response to the input control word, the output instantly steps in 19.5kHz increments over 0 to 13MHz. The DDFS ROM is 32 times smaller than in a straightforward design. Half- and quarter-wave symmetry of the sine function are exploited to generate sine and cosine waveforms from one ROM by phase shift of the argument. Internal word length is shortened by 2b because the ROM stores the difference between sine amplitude and phase. One large lookup table is replaced by smaller coarse and fine-interpolation tables.

The DAC successively bisects a reference charge between two equal capacitors through a switch controlled by the digital input, and therefore dissipates no static power. Its dynamic CV^2f dissipation gets smaller with capacitance C, whose minimum value is bounded by mismatch errors and thermal noise. At the end of conversion, a sample-and-hold buffer translates charge to voltage. If the buffer is not slew-rate limited, resampling will guarantee that no code-dependent glitches appear in the output. A conversion rate of 50MHz is obtained by pipelined operation of ten charge-redistribution stages. Increasingly significant bits of the data word are applied through delay registers to ten one-bit charge-redistribution cells in cascade, each cell consisting of one capacitor and two switches (Figure 2) [2]. Each data bit initially precharges its cell capacitor to a high or low reference voltage. When the capacitor is shorted to the previous cell, the resulting average voltage corresponds to a 2b D/A conversion. A 10b conversion is obtained after the charge traverses the ten cells. As in a CCD, the pipelined charge-transfer requires a three-phase non-overlapping clock.

*Research supported by ARPA, Rockwell International, and the State of California MICRO Program.

The linearity in this binary-weighted DAC is limited by capacitor mismatch, and signal-dependent stray charge injected by the MOSFET switches. The conversion rate is limited by the RC settling time in each cell. If FETs used to reduce the RC time constant are too wide, their voltage-dependent junction capacitance in parallel with the cell capacitor will add nonlinearity. Using published data on capacitor mismatch and known FET models, the cell components are optimized within these constraints, leading to a 500fF cell capacitor and a complementary switch composed of 5μm-wide nMOS and pMOS FETs. With 3V drive at FET gates, an acceptably short RC time constant is obtained when swings on the capacitor are limited to ±0.25V about a 1V common-mode level (Figure 3).

Three cells at a time are precharged by the respective data bits, followed by redistribution of charge among the cells under control of the three-phase clock. Accordingly, the input word is divided into three 3b nibbles and one MSB, and after appropriate delay in the pipeline registers, each nibble drives a 3b DAC stage. The equal-sized FETs in the complementary switch to the first-order cancel each other's charge injection. Stray charge injected by the switch at the conclusion of precharge causes a DAC gain error. Injection after charge-sharing is signal-dependent and contributes nonlinearity. This source of error is acceptably small for 10b operation.

The DAC output charge is converted by a buffer into a voltage. To attain a linearity commensurate with the DAC, the buffer is a balanced switched-capacitor unity-gain stage using a super-cascode op amp with about 60dB open-loop gain (Figure 2) [3]. Two DAC pipelines driven by complementary data apply a balanced input to the buffer, whose clocking scheme cancels the signal-dependent charge from its own sampling switches. An open-drain degenerated differential pair in the op amp feedback loop drives off-chip instruments at 50Ω. The last stage in the charge redistribution pipeline, shown driven by a data 1, boosts the common-mode voltage on the capacitors from 1V in the pipeline to 1.5V at the buffer input. The op amp sampling switches and capacitor are configured so that the buffer input itself resembles a DAC cell. The three clock phases in the DAC and the two additional phases for the buffer are derived on-chip from a single external clock (Figure 4).

At 40 MHz with 3V supply, the DDFS logic dissipates 35mW, the DAC clock generator, 4mW, and the DAC charge redistribution pipeline, less than 1mW. In this prototype a large output buffer drives off-chip loads. When scaled to drive typical on-chip capacitive loads, it is expected to dissipate 3 to 5mW. The worst-case spurious products in a DDFS system operating at a sample rate fs appear in sinewaves synthesized at frequencies fs/4 and fs/3. Spurious frequencies are below -72dBc. At low output frequencies, 3rd harmonic is -56dBc while in an output tone at a slight offset from fs/3 the largest aliased harmonic is at -50dBc (Figures 5a and 5b). These harmonics are higher than expected, and arise from a small inter-wire fringing capacitance linking DAC cells (Figure 6). Results from an IC with a revised layout with this parasitic reduced are expected to give close to 10b linearity.

References

[1] Nicholas, H.T., H. Samueli, "A 150MHz Direct Digital Frequency Synthesizer in 1.25μm CMOS with -90dBc Spurious Response," IEEE J. of Solid State Circuits, v. 26, pp. 1959-1969, Dec., 1991.

[2] Wang, F-J., et al., "A Quasi-Passive CMOS Pipeline D/A Converter," IEEE J. of Solid State Circuits, v.24, pp. 1752-1756, Dec., 1989.

[3] Bult, K., G. Geelen, "A Fast-Settling CMOS Op Amp for SC Circuits with 90dB dc Gain," IEEE J. of Solid State Circuits, v.25, pp. 628-636, Dec., 1990.

0-7803-1844-7 / 94 / $3.00 / © 1994 IEEE

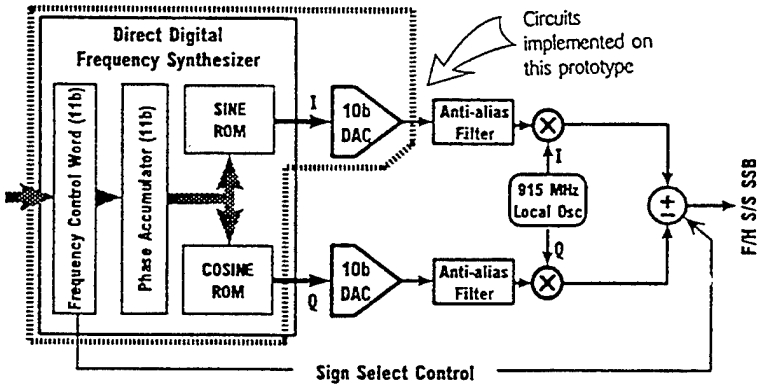


Figure 1: Block diagram of frequency synthesizer.

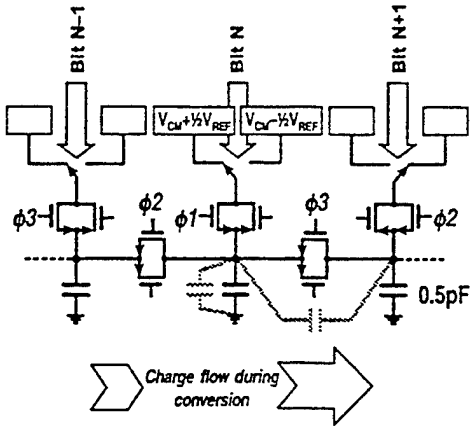


Figure 3: Charge-redistribution cell. Parasitic capacitance between cell capacitors (gray) causes integral nonlinearity.

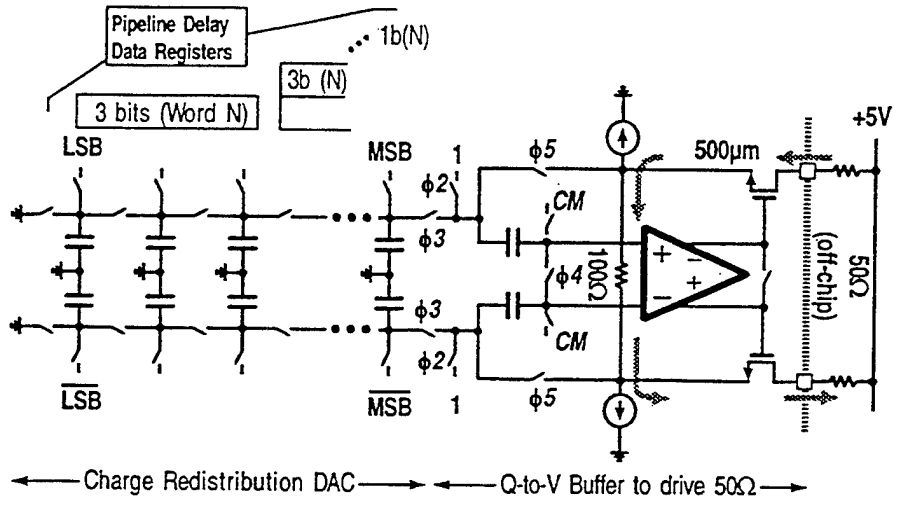


Figure 2: D/A converter uses pipelined charge redistribution in balanced configuration.

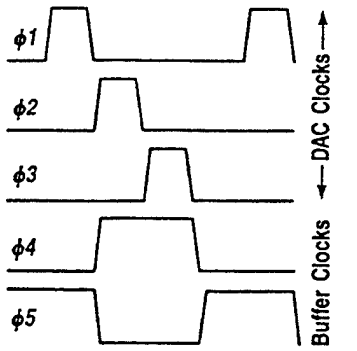


Figure 4: Clock phase required in D/A converter, buffer.

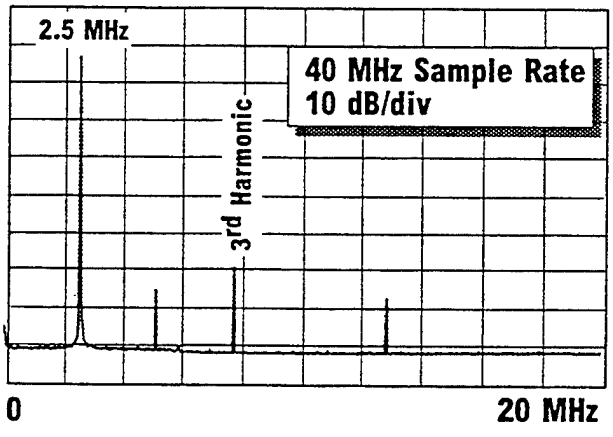


Figure 5a: 2.5MHz (low freq.) output sine spectrum.

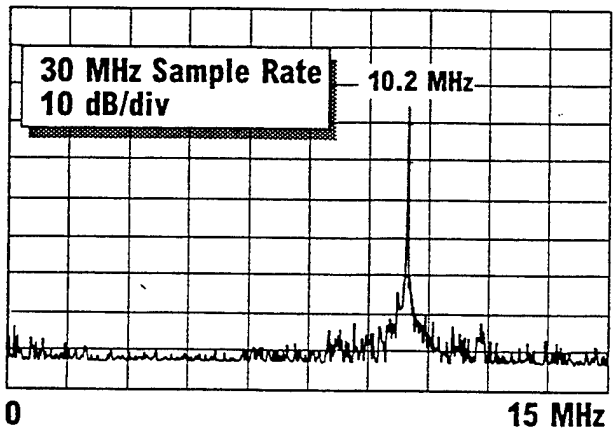


Figure 5b: 10.2MHz sinewave spectrum (fs/3 offset).
Figure 6: See page 304.

WP 2.2: A Low-Power CMOS Digitally Synthesized 0-13MHz Agile Sinewave Generator*

Glenn Chang, Ahmadreza Rofougaran, Mong-Kai Ku, Asad A. Abidi, Henry Samueli

Electrical Engineering Department, University of California, Los Angeles, CA

G. Chang

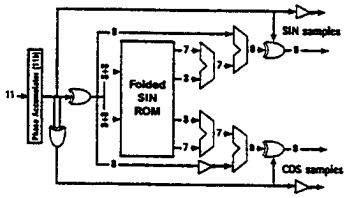
A 40mW direct digital frequency synthesizer for frequency hopping spread spectrum communication having a 10b D/A converter produces output sinewaves whose frequency is selected with an 11b input word. The 2.9x4.9mm² IC in 1μm CMOS operates with a 40MHz clock at 3V.

See Digest page 32.

Outline

- ☐ Agile Frequency Synthesis in Wireless Transceivers
- ☐ Direct Digital Frequency Synthesizer
- ☐ D/A Converter
- ☐ Measurements of Spectral Purity
- ☐ Conclusions

Direct Digital Frequency Synthesizer

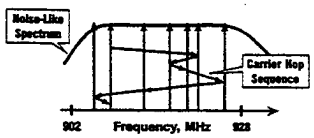


- DDFS guarantees spurious levels < −72 dBc
- Output frequency resolution is (Sample Rate)/2¹¹
  - ✓ ROM contains only quarter-wave data
  - ✓ SIN and COS generated from same ROM by phase-shift of argument
  - ✓ ROM stores difference between amplitude and phase (saves 2 bits)
  - ✓ One large table is replaced by small coarse and fine tables

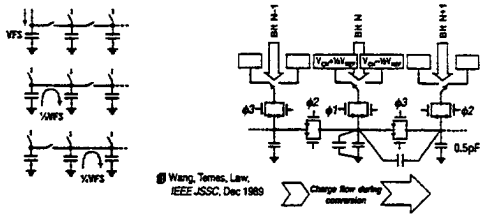
⇨ ROM is 32× smaller

Frequency-Hopped Spread Spectrum

- Spread-spectrum communications allow a large number of users can share the same spectrum; intended user searches for particular spreading code
- FCC has allocated 902-928 MHz band for unlicensed, spread-spectrum use
- Techniques of spectrum spreading: *direct-sequence* or *frequency-hopping*
- Frequency-hopped spread-spectrum allows wideband spreading at any data rate (⇨ low power dissipation), but needs **agile frequency source**



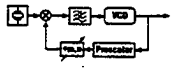
Principle of Low-Power, High-Speed DAC



- Binary division by successive charge redistribution
- Equal-sized capacitors required
- Three-phase clock for proper charge-transfer
- Pipelined operation produces one conversion per clock cycle
- Linearity limited by:
  - ⇨ DAC capacitor mismatch
  - ⇨ Stray capacitance in DAC cells
  - ⇨ Signal-dependent charge injection after redistribution

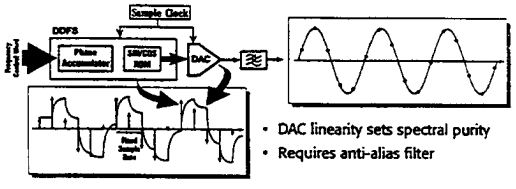
Methods of Frequency Synthesis

Phase-Locked Loop



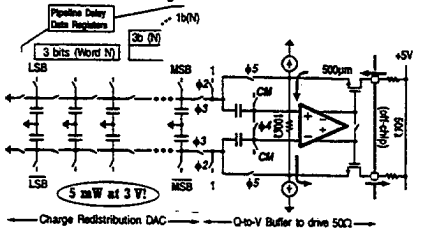
- Frequency agility limited by loop settling time
- Signal purity and wide tuning range compromise VCO design

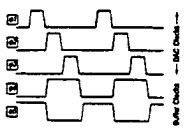
Digital Synthesizer & D/A Converter



- DAC linearity sets spectral purity
- Requires anti-alias filter

DAC Implementation





- Differential implementation using two charge-redistribution pipelines
- Output buffer must be at least as linear as DAC – differential buffer degenerated by polysilicon resistor, and driven in closed-loop

What Sets DAC Power Dissipation?

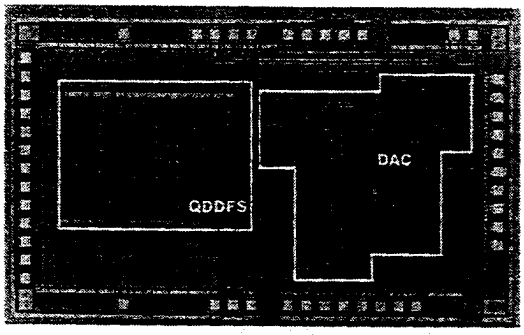
Sources of Power Dissipation

- No static power dissipation in DAC core, but small dynamic CV^2f dissipation
- Clock buffers driving DAC switches dissipate most power

  ↻ Power dissipation decreases as DAC cells are scaled down

Lower Limits to Scaling

- 0.5 pF capacitors (400 sq-μm area) *match* to within 0.1% rms
  Pelgrom, *et al.*, IEEE JSSC, Oct 1989
- Switch-induced *noise* with 0.5 pF capacitors accumulating in DAC ≈ 170 μV rms; output buffer *noise* ≈ 110 μV rms ↻ LSB size > 0.5 mV
- *RC time constant* for settling to 10 b at 50 MHz sets width of switch FET ↻ lower limit on nonlinear charge injection

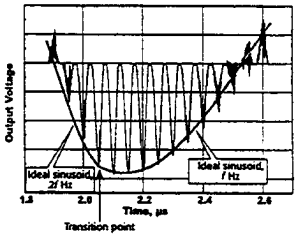


1-μm CMOS with double-metal, linear capacitor
2.9 × 4.9 mm die size

Conclusions

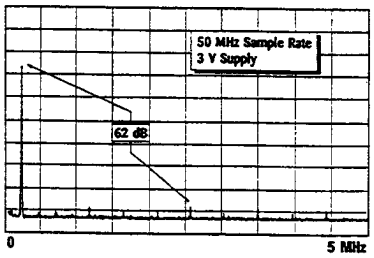
- ✓ First demonstration of monolithic CMOS 10b DDFS/DAC
- ✓ Low-power design leads to 35 mW DDFS, and 5 mW 10b DAC core, both operating at 50 MHz from 3-V
- ✓ Spectral purity from untrimmed parts is –62 dBc at low frequencies, –57 dBc at 1/3 Sample Rate
- ✓ Low-power circuits ↻ small interaction between analog and digital parts of the chip
- ✓ Direct digital frequency synthesis is a viable solution for an agile sinewave source in battery-powered wireless transceivers

Frequency Agility



- Output instantly switches from one frequency to another (after 8 clock cycle latency through DDFS/DAC)
- Anti-alias filter will limit the settling response in system

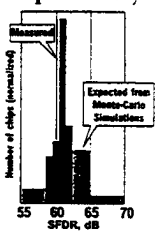
Low-Frequency Synthesis



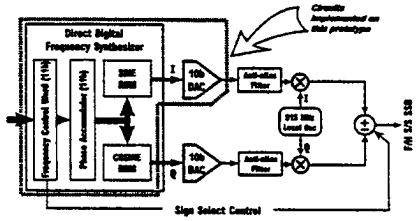
✓ Noise floor set by quantization noise
– measure 2 dB higher than theoretical limit

✓ Spurious level as predicted
– set by capacitor mismatch

Repeatability

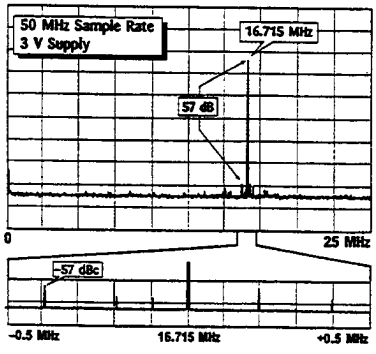


Frequency-Hopping RF Transmitter



- Single-step I-Q upconversion produces single-sideband, suppressed-carrier output in the 902-928 MHz band
- DDFS/DAC need only span 0→13 MHz: sign-select at output produces 902-915 MHz, or 915-928 MHz
- Order of anti-alias filter depends on highest output frequency relative to sample rate
- Acceptable image suppression requires 8b matching in two channels

High-Frequency Synthesis



- Spurious levels grow at high frequencies due to 1fF inter-cell stray capacitance
- *No* slew-rate limiting or output glitch at 50 MHz
- On-chip digital circuits do *not* contaminate output spectrum!

Channel Coding for a Frequency-Hopped Wireless Transceiver

V.S. Lin and G.J. Pottie

Lecture Notes in Computer Science
A.T. Goldberg and N.P. Secord, eds.
pp. 109-124, Springer-Verlag, 1994

Channel Coding for a Frequency-Hopped Wireless Transceiver[1]

Victor S. Lin and Gregory J. Pottie

Electrical Engineering Department
University of California, Los Angeles
405 Hilgard Ave.
Los Angeles, CA 90024

Abstract - We consider a multiple-access frequency-hopped system employing antenna diversity and error correction coding. A construction based on orthogonal Latin squares is used for the hopping patterns to improve code performance in the presence of interference. The measures of interest are the system performance, delay and decoder complexity, providing a basis for comparing BCH, Reed-Solomon (RS), and convolutional codes. It is shown that when binary non-coherent FSK modulation and hard decision detection are used, the BCH, RS, and convolutional codes that can meet the delay and bandwidth constraints have similar performance and comparable complexity. Soft decision decoding for the convolutional code and error-and-erasure correction decoding for the RS code are investigated in terms of the performance-complexity trade-offs for different decoding metrics. It is shown that improvement over hard decision decoding is significant when the right decoding metric is selected.

1. Introduction

The dominant impairments to transmission in cellular radio systems are multipath fading and interference from other users sharing the common spectrum. Multipath propagation results in severe fading, which can dramatically increase the signal to noise ratio (SNR) required for reliable operation. In a time-division multiple access (TDMA) system the users are assigned slots which are kept from frame to frame, with interfering users in nearby cells assigned slots in the same way. Because of factors such as geographic proximity and shadow fading, the result is that interference power also displays wide variations across slots, which further degrades the 90% to 99% worst case performance. The conventional solution of restricting the frequency re-use between cells significantly reduces capacity. Since the re-use distance must be made large enough to allow sufficient margin against multipath fading, a high performance price is paid.

In a direct sequence CDMA system, the problem of differing interference levels is mitigated by the despreading process. Effectively, for large enough spreading each

1. This work supported in part by AT&T Bell Laboratories Mathematical Sciences Research Center and ARPA contract DAAB07-92-R-C977.

user deals with the average interference levels, and thus the system may be engineered to deal with the average rather then worst case interference levels. Channel coding completes the task, and further reduces the SNR required for reliable operation. Unfortunately, wideband direct sequence systems suffer from two problems: (1) the power consumption of the receiver is very large and (2) synchronous access is not possible on the up-link, since the chip duration is very short. The latter problem results in decreased capacity, since the interference generated within the cell considerably exceeds the interference generated in neighboring cells. A frequency hopped system can alleviate both of these difficulties since the hopping rate is independent of the bandwidth occupied, and therefore can be set low enough to permit low-power receiver design and synchronous operation.

An ongoing research project at UCLA is the design of a low-power hand-held transceiver [1]. Data rates between 2 kb/s to 160 kb/s are to be supported, using a 26 MHz bandwidth. The maximum hop rate is 160 khops/s and the maximum range is 500 meters. Experimental characterization of the radio channel has found that the typical delay spread is less than 400 ns for both indoor and outdoor microcellular environments [2][3]. At the maximum hop rate, the hopping or baud interval (6.25 μs) is large relative to the delay spread; hence, the effect of ISI caused by multipath is negligible and equalization is not needed. In order to achieve a high capacity with low power consumption, it is essential that diversity be employed. Dual antenna diversity, frequency hopping, and channel coding will be implemented. The frequency hopped transceiver will consume less than 300 mW in active mode with two receiver branches, in comparison to well over 1 Watt of power for a DS system implemented in the same technology [4]. This is possible because of a careful design of the digital direct frequency synthesizer/digital-to-analog converter pairs [5]. In the following, we outline our progress in the design of the hopping patterns and selection of the channel codes, and the basic trade-offs in choice of metrics for decoding.

2. Latin Squares

Slow frequency hopping in combination with interleaving and channel coding is effective in dealing with variations in the signal to interference ratio among successive hops, provided the hopping patterns are selected properly. A number of recent papers [6][7][8] have examined slow-hopped systems. We will concentrate a simple construction introduced in [9] which is especially suitable for this application.

A latin square of order n is an n x n matrix with entries from a set R of n distinct elements, say R = {0,1,...,n-1} such that each row and column contains every element of R exactly once. Two Latin squares are said to be orthogonal if the n^2 ordered pairs (i,j), where i and j are the entries from the same position in the respective squares, exhaust the n^2 possibilities, i.e., every ordered pair occurs exactly once.

If n is prime, then there is a simple construction for a family of n-1 mutually orthogonal latin squares. For a = 1,...,n-1 we define an n x n matrix {a} by setting

$$\{a\}_{ij} \equiv ai + j \pmod{n}$$

It is easy to verify that each matrix is a latin square: for fixed i, the numbers ai+j,

j=0,...,n-1 are distinct, and for fixed j, the numbers ai+j, i=0,...,n-1 are distinct, in both cases assuming modulo arithmetic. It remains to verify that if $a \neq b$, then the matrices {a} and {b} are orthogonal. This amounts to showing that for all c, d = 0,1,...,n-1, there is a unique solution (i,j) to the linear system

$$c \equiv ai + j$$
$$d \equiv bi + j$$

This follows from the fact that the determinant of this linear system is a-b, which is non-zero (mod n).

If n is prime, then the residues modulo n form the finite field of order n. The above construction extends to finite fields of prime power order in the natural way: now R is the set of field elements, and for every non-zero field element a, we define a matrix {a} with the ij^{th} entry $\{a\}_{ij} = ai+j$, where addition and multiplication take place in the field.

The above construction may be used in a synchronous cellular radio system. That is, the timing of the slot boundaries must be aligned across cells, with guard bands potentially required for large cells to account for propagation delays from nearby cells. Each cell is assigned a matrix {a} in the set of n-1 mutually orthogonal latin squares. Then up to n users can be accommodated in any cell, each user being assigned a field element. For example, suppose n=8, a=5, and a user is assigned the element 7. We might interpret the rows as frequency slots and the columns as time slots for a frequency hopped system. Then at time 0, frequency 3 is transmitted, at time 1 frequency 6 is transmitted, etc. Except for other cells which also are assigned matrix {5} there will be at most one collision per user at any given time/frequency pair over the 8 slot duration of the frame. If a regular re-use pattern of 7 cells is used for the matrices, then the signals from users sharing the same hopping patterns will be heavily attenuated. In this way, each user suffers collisions from an ensemble of users occupying the nearby cells, rather than with one dominant interferer. Obviously, with a larger value of n, the re-use distance increases and the randomization of interference hits is more complete. Also, the rows might be reinterpreted as time slots in a frame, and the columns as frame positions within a superframe for a pure TDMA system.

The construction allows each mobile to quickly generate its own time/frequency hopping pattern given minimal information transfer from the base station. The only information required, given that n and the primitive polynomial p(x) are known, is the matrix number, a, and the field element, e. Then if i is the row number, the column corresponding to i is just $j \equiv e + ai \pmod{p(x)}$. Hence, the overhead involved in communicating the hopping pattern is negligible. Note that we cannot get every possible value of n using this construction. However, primes and prime powers are dense for moderate n, affording a wide range of choices.

A very similar construction has recently been proposed by Wyner [10] using a combinatorial configuration constructed by Roche [11] for use in wideband multi-tone systems, to assure that users in nearby cells interfere in at most one tone.

The effectiveness of the Latin squares construction was illustrated in [9] for a slow-hopped DQPSK system. Simulations were conducted in combination with rate-1/2 and rate-1/4 convolutional codes with 32 states, for a multi-cell system using the

same propagation assumptions as in [12]. It was further assumed that Rayleigh fading was independent between frequency hops, and that channel parameters were estimated from the data. The results indicate that the system capacity without some randomization of interferers is quite poor, but with the Latin squares approach capacity increased beyond that reported for a DS-CDMA system in [12], with considerably lower decoding delay. The addition of dual antenna diversity further improved performance.

Simulations were also conducted to examine performance in flat fading conditions. The results in this case are particularly interesting in illustrating what we refer to as interferer diversity. Even though the desired signal level is now at the same level in each slot, the interference levels are independent, since a different set of interferers in encountered in each slot. Thus, even in flat-fading conditions, the interference is independently fading, leading to independent signal to interference ratios. Therefore channel coding continues to provide a significant improvement in performance, effectively averaging over the SNR variations while providing coding gain. This is of considerable interest in handheld transceiver applications, since fading can be flat over tens of MHz in some indoor situations. In these cases, frequency hopping in itself would not be sufficient to achieve diversity benefit; some randomization of interference is also required. Consequently, the Latin squares construction will be incorporated in the final design.

3. Channel Code Selection

Fig. 1 outlines the system model used in the investigation of alternative channel coding schemes. The input bit stream is encoded into the coded q-bit symbol sequence of a block code or a convolutional code and the resulting coded sequence with a rate of f_c symbols/s is fed to the interleaver. The interleaver rearranges the coded sequence in such a way that the fading channel with burst error characteristic is transformed into a channel having independent errors. The coded symbol at the output of the interleaver, to be transmitted with a rate of $f_c \times q$, is assigned to one of the two binary orthogonal FSK signals.

The frequency hopping binary FSK signal is output from a frequency hopper where the hopping local signal is multiplied to the binary FSK signal. A slow frequency hop scheme is assumed in which a hop is made in each slot interval consisting of a few symbols. It is assumed that the received signal experiences independent Rayleigh fading slot-by-slot, and that it is perturbed by additive white Gaussian noise. While this model does not yield a capacity estimate, it is sufficient to evaluate the relative effectiveness of coding schemes in a situation where we in fact expect the SNR to be independent from hop to hop as a result of the use of the Latin squares construction.

The receiver dehops the signal to obtain the received binary FSK signal. Noncoherent detection is used for the reception, because the handheld system is designed to allow fast hopping. The receiver employs dual antenna diversity with postdetection equal gain combining to achieve diversity. This was found to outperform selection diversity, and in any case when the channel SNR cannot be estimated with any accuracy for a slot (e.g., fast hopping), two complete receiver branches are required to make antenna diversity effective. Thus, equal gain combining comes essentially for

free. We assume that the signals received at the two antennas have low correlation.

As a result of interleaving/deinterleaving and frequency hopping, errors within a code word appear to be independent. We assume that a convolutional interleaver is used because it results in less delay than block interleaving. In this section, we consider hard decision decoding for convolutional codes and error correction for block codes. Later, we examine soft decision decoding for convolutional codes and error-and-erasure correction for block codes.

In order to compare the various coding options in terms of BER performance, we assume ideal interleaving so that the channel is memoryless. For non-coherent binary FSK (NC-BFSK) modulation with ideal interleaving, the error rate performance over a frequency-nonselective, slowly fading channel is

$$P_2 = \frac{1}{2 + \bar{\gamma}_b} \tag{1}$$

where is $\bar{\gamma}_b$ the average signal-to-noise ratio, defined as

$$\bar{\gamma}_b = \frac{\xi_b}{N_o} E(\alpha^2) \tag{2}$$

The term $E(\alpha^2)$ is the average value of the Rayleigh distributed envelope squared.

When L-th order antenna diversity technique is used, the performance of square-law-detected binary FSK is well approximated by [13]

$$P_2 \approx \left(\frac{1}{\bar{\gamma}_b}\right)^L \binom{2L-1}{L} \tag{3}$$

For dual antenna diversity, L is set equal to two.

3.1 Performance of Reed-Solomon and BCH Codes

With Reed-Solomon (RS) codes over GF(2^q), q binary channel symbols are used to form one code symbol. Therefore, the probability of code symbol error is $P = 1 - (1 - P_2)^q$, where P_2 is given by (1) and (2), with and without antenna diversity, respectively. For M-ary ($M = 2^q$) block codes with error correction decoding, the decoded symbol error probability can be approximated as

$$P_{es} \approx \frac{1}{N} \sum_{i=t+1}^{N} i \binom{N}{i} P^i (1-P)^{N-i} \tag{4}$$

where $t = [(d - 1)/2]$ is the number of errors that can be corrected by the code, d is the minimum distance of the code, N is the block length, and [x] is the largest integer less than or equal to x. The probability of a decoded bit error is

$$P_{eb} = \frac{2^{q-1}}{2^q - 1} P_{es} \tag{5}$$

For BCH codes, the probability of decoded bit error is just $P_{eb} = P_{es}$, where P_{es} is given by (4). Thus, by combining (4), (5) and (2) or (3), the probability of decoded bit error for block codes can be computed at any signal-to-noise ratio, with or without antenna diversity.

3.2 Performance of Convolutional Codes

Next, we consider the error rate performance of convolutional codes decoded using the Viterbi algorithm with a hard decision decoding metric. The decoded bit error probability of a rate b / n convolutional code can be upperbounded by

$$P_b < \frac{1}{b} \sum_{l=d}^{\infty} \beta_l D^l \tag{6}$$

where d is the minimum free distance of the code, and β_l is the total information weight of all paths of distance l from the all-zeros path. With hard decision decoding

$$D = \sqrt{4P_2(1-P_2)} \tag{7}$$

where P_2 is given by (1) and (2), with and without antenna diversity, respectively.

The weight and distance structure of many convolutional codes have been computed and tabulated in the literature. Thus, equations (6) and (7) can be used to evaluate the performance of a specific convolutional code.

To study the performance advantage of coded systems over an uncoded one, a system simulation was designed and completed. For decoding Reed-Solomon and BCH codes, the Berlekemp-Massey algorithm and Forney algorithm were implemented. The Viterbi algorithm was employed for decoding convolutional codes.

The simulation results were checked against the corresponding analytical results whenever possible in order to verify the accuracy of our simulation software during the initial phases of the simulation development. It was found that the assumption of ideal interleaving was good. Analytical expressions (5) and (6) produced fairly tight upper bounds. For each coding scheme, the difference between the performance curves obtained from the upper bound and the simulation is less than 1 dB for $10^{-2} < \text{BER} < 10^{-4}$. The simulation results were used to compute coding gains. Fig. 2 illustrates typical BER curves generated from the simulation data for a particular convolutional code with and without dual antenna diversity. Some of the coding gain data are summarized in Table 1. It shows that using the rate-1/2, 32 state (v=5) convolutional code resulted in near minimum SNR at the target BER, with and without antenna diversity. The coding gain achieved is 19.1 dB with antenna diversity, compared to 13.0 dB with a single antenna. This highly motivates the use of both channel coding and antenna diversity in the system architecture.

The code rate and code size for the codes in Table 1 were chosen based on their

potential to meet code performance with decoding delay tolerable for voice transmission, practical code complexity, and channel bandwidth constraints.

3.3 Decoding Delay

To evaluate and compare the delay of the selected codes, we defined three different delay measurements: decoding delay, interleaving delay and overall delay. The delay quantities are specified in signaling interval units, Ts, to provide a normalized delay representation. Denoted τ_d, decoding delay is the waiting time for some number of encoded symbol to be received before starting the decoding process. The interleaving delay, τ_i, for a (I, J) convolutional interleaver is proportional to $(I-1) \times J$, where values of I and J are chosen to randomize the error bursts of the fading channel and depend on the coding scheme, the code parameter, and the number of channel symbols per frequency slot [14]. The overall delay is the sum of the decoding delay and interleaving delay, i.e. $\tau_c = \tau_d + \tau_i$.

To show how delay is determined by the system architecture, examine a system with BFSK modulation and a Reed-Solomon code over $GF(2^q)$. The code has a block length spanning $(2^q - 1) \times q$ channel symbols and thus, $\tau_d/Ts = (2^q - 1) \times q$. Assume there are $c \times q$ channel symbols per slot or equivalently, c code symbols per frequency slot. In order to ensure that all code symbols in a code word suffer from independent fading, successive code symbols must be interleaved across slots. This requires that $I = c$ and $J = (2^q - 1)$. Therefore, $\tau_i/Ts = (c - 1) \times (2^q - 1) \times q$ and $\tau_c/Ts = (2^q - 1) \times c \times q$.

Now, consider a BFSK system with c x q channel symbols per frequency slot. In this example the coding scheme is a rate-1/2 convolutional code with memory v (no. of states = 2^v). It has been shown that the full minimum distances of the 32 the 64 state codes are obtained with a truncation depth of 19 and 27, respectively [15]; these parameters were used in computing the delays given in Table 1. For the purpose of delay comparison, we assume that the full minimum distance is obtained with a truncation depth of $4.5 \times v$. This implies $\tau_d/Ts = 9 \times v$. Setting $I = c \times q$ and $J = 9 \times v$ as the interleaving parameters guarantees independent fading condition for the successive symbols generated by the encoder with a separation less than truncation depth. Thus, $\tau_i/Ts = (c \times q - 1) \times 9 \times v$, and $\tau_c/Ts = 9 \times v \times c \times q$.

It can also be shown that using BCH codes with block length N and bit-by-bit interleaving result in $\tau_d/Ts = N$, $\tau_i/Ts = N \times (c \times q - 1)$, and $\tau_c/Ts = N \times c \times q$. Note that for similar system parameters, the overall delays for the BCH and Reed-Solomon codes with equal block lengths are the same. The overall delay expressions derived above imply that a Reed-Solomon code over $GF(2^q)$ has comparable overall delay to a convolutional code with memory $v = 2^q/9$.

Codes satisfying the voice transmission delay requirement are listed in Table 1. To compute the delay in seconds, we assumed that the delay constraint is 20 ms, each slot consists of 6 BFSK signals for the length 63 block codes and 10 for remaining codes, and the channel transmission rate is 16 kbit/s.

Table 1: Code Performance and Delay Comparison

Code Type RS = Reed-Solomon CC = Convolutional	Gain w/Single Antenna (dB)	Gain w/Dual Antenna (dB)	Coding Delay w/o Interleaving (Ts)	Overall Coding Delay (Ts)
BCH (n=31, k=16, d=7)	11.5	17.9	31	310
BCH (n=63, k=36, d=11)	12.3	18.3	63	378
RS (n=31, k=15, d=17)	13.1	18.2	155	310
RS (n=63, k=33, d=31)	13.2	18.7	378	378
CC (r=1/2, v=5)	13.5	19.3	38	380
CC (r=1/2, v=6)	13.6	19.5	54	540

To ensure practical code complexity, code selections were limited to block codes with block length less than 255 and convolutional codes with 128 states or less. These are codes which are currently used in various digital communication systems. Channel bandwidth constraints were included in the code selections by limiting the code rate to approximately 1/2; this limits the channel bit rate to a value no larger than twice the user data rate. This limitation was imposed to avoid excessively large bandwidth requirements with high data rates and to decrease link vulnerability from frequency selective fading at high data rates.

3.4 Decoder Complexity

Code complexity was quantified in terms of multiplies per decoded symbol for the block codes and additions per decoded symbol for the convolutional code. We found that decoding a single t-error correcting BCH code word required approximately $10t^2 + 3tn$ multiplications, obtained by adding $10t^2$ multiplications for executing the Berlekemp-Massey algorithm to $2tn$ for evaluating the syndrome, and tn for Chien search. For a t-error correcting RS code, the number of multiplications required to decode a code word increases due to the multiplications in the Forney algorithm. However, the number of operations normalized per bit for RS codes is not necessarily higher than BCH codes when the code length and rate are fixed. In fact, our calculation showed that the (63, 33) RS codes required 26 multiplies per bit versus 33 multiplies per bit for the (63, 36) BCH code. For a 2^v convolutional code, a total of $3*2^v$ addition operations are necessary at each stage. This includes two additions to compute the cumulative path metrics of the paths merging at each state and a comparison (subtraction) to determine which incoming path survives. For convolutional codes, in addition, a trace back operation is required to complete the decoding. For the codes parameters listed in Table 1, the convolutional codes required more than twice the operations per bit than the comparable block codes. This fact was partly confirmed by the longer simulation times for convolutional codes.

The results of analysis and simulation indicate that for a system architecture employing slow frequency hop and dual antenna diversity to combat multipath fading, a BCH code performs as well as an RS code with a comparable code rate. Our investigation also shows that a rate-1/2, 32 state convolutional code can attain a relatively large coding gain at 10^{-3} BER while meeting the tight delay constraint for two-way speech transmission. In addition, convolutional codes offer the advantage of efficient soft decision decoding. The conclusion is that complexity and performance for the code candidates are similar for hard decision decoding; however, the dependence on soft decision decoding must be determined before deciding which code is the best candidate. We now examine some of the relevant trade-offs.

4. Decoding Metrics

For Rayleigh fading channels, soft decision decoding with perfect channel parameters can effectively double the diversity order available through coding [13]. However, for a real system where channel parameters are obtained by an imperfect estimator, the unreliable estimates could significantly degrade the performance benefits of soft decoding with side information. The reliability of the channel parameter estimator, the modulation scheme, and the channel condition are some of the factors that affect the formulation of a good soft decoding metric. Furthermore, the trade-offs in performance and complexity should be considered when selecting a decision metric for NC-BFSK systems because most soft decoding schemes require much higher receiver/decoder complexity than hard or erasure decoding.

4.1 Soft Decision Metrics for Fading AWGN Channels

The optimum soft decision metric is derived from the likelihood function of the decision variables. For a NC-BFSK system, in which one of two frequencies is transmitted with equal probability to a receiver and the transmitted signal is corrupted by additive white Gaussian noise with spectral density $N_o/2$, a model of the received signal is

$$r_i(t) = \alpha \sin(\omega_i t + \phi) + n(t); \qquad i = 0, 1 \tag{8}$$

We assume that the amplitude, α, is known and the phase, ϕ, is a random variable uniformly distributed in the interval $[0, 2\pi]$. It can be shown that when the constant terms are eliminated, the log-likelihood functions may be written

$$M_{ML} = \ln(p_i(r)) \sim \ln\left(I_o\left(\frac{2\alpha z_i}{N_o}\right)\right); \qquad i = 0, 1 \tag{9}$$

where z_i is the decision variable produced by a square law detector (the optimum demodulator for NC- FSK), and $I_o(\cdot)$ is the modified Bessel function [16].

The mapping of the decision variables into the maximum-likelihood (ML) soft metric involves a very complicated function, $\ln(I_o(\cdot))$. To implement the branch computation part of the Viterbi decoder, a look-up table will be required to transform the

decision variables into branch metrics. Since the transforming function has a linear and a non-linear region, one way to reduce the size of the look-up table is by storing the values of the function over the non-linear region and use a linear approximation formula, which does not require costly memory storage, to compute the function over the linear region. Using this procedure, the size of the look-up table used in our study is about 2 Kilobytes.

Besides the ML decoding metric, we consider two suboptimum but less complicated soft decision metrics. The first is based on the Euclidean distance concept. The Euclidean distance metric has been shown to be the optimal soft metric for coherent PSK systems in AWGN channels, but it is not optimal for NC-BFSK systems. It may be written:

$$M_E = (z_i - \alpha)^2 + (z_j)^2 \tag{10}$$

where $i = 0$ and $j = 1$ for the hypothesis that z_0 was received, and vise versa for the hypothesis that z_1 was received. The metric is simpler not only because it does not involve any complicated functions but also the fade magnitude of the signal is the only channel parameter appearing in the metric. The other metric we considered required even less processing and complexity by using a simple linear combining scheme:

$$M_{LC} = z_i - z_j \tag{11}$$

where values of i and j are defined as in the previous metric. The linear combining metric does not require any channel state information.

4.2 Soft Decision Metrics for Multiple-Access Channels

FH/SS systems operating in a multiple-access channel will experience background thermal noise, as well as, interference from other users. We use the following model of the received signal for the multiple-access channel:

$$r_i(t) = \alpha \sin(\omega_i t + \phi) + \beta i(t) + n(t); \qquad i = 0, 1 \tag{12}$$

where the fade levels α and β are independent Rayleigh distributed random variables $i(t)$ and $n(t)$ are independent narrowband white Gaussian processes. Since interference is typically the dominant impairment in multiple-access environments, we can assume that the noise term is negligible and set the two-sided spectral density of $i(t)$ to $N_o/2$. Thus, when $E[\beta^2]$ is normalized to one, then

$$E[\beta i(t)\beta i(\tau)] = E[\beta^2]\,E[i(t)i(\tau)] = \frac{N_o}{2}\delta(t-\tau) \tag{13}$$

Note that for the FH system under consideration, the only real difference between the fading AWGN channel and the multiple-access channel is that the power of the "noise" in the latter is changing from hop to hop.

To perform well for the multiple-access channels, metrics (9) and (10) were reformulated to account for the variation in noise power in each symbol interval. The re-

formulations are given by (14) and (15), respectively:

$$M_{L,1} = \ln\left[Io\left(\frac{2\alpha z_i}{\beta^2 N_o}\right)\right] \tag{14}$$

$$M_{E,1} = \frac{(z_i - \alpha)^2 + (z_j)^2}{\beta^2 N_o} \tag{15}$$

For NC-BFSK, the relevant quantities for soft decision decoding are the decision variables for the two frequencies, the received signal power, and the noise power for each slot. We next present ways of obtaining the channel parameters for slow-hopped systems.

4.3 Estimation of Channel Parameters

The estimation of channel parameters from the data and with training sequences was investigated in [9], in the context of DQPSK. It was found that it is better from the point of view of capacity to form estimates directly from the data-bearing signals for short slots of 8-16 symbols, rather than appending a training sequence. In addition, it was found that for some combinations of cell loading and channel codes it was better to use an erasure-declaring mechanism than to use the soft metric proposed; but in any case performance was always better than using simple hard decisions for a properly chosen erasure threshold. Simulations were also conducted using the soft metric with perfect channel knowledge, revealing a very large gap in performance. Thus, channel state information can be very valuable in decoding.

For NC-BFSK signaling with square-law detection, the signal power for a slot can be estimated by accumulating the larger decision variable, z_i, for each received symbol in the slot. This type of estimation involves hard decision demodulation. While it is relatively easy to estimate the power of the desired signal for useful signal to interference ratios, it is more difficult to accurately estimate the interference power. For NC-BFSK, an orthogonal signaling scheme, one way to form an estimate of the noise (or interference) power for a slot is to accumulate the smaller decision variable, z_j, in the slot. The reliability of the signal power and noise power estimates depends on the sample size of the estimator, which is equal to the number of symbols per slot.

4.4 Erasure Metric

When the channel state information are not so reliable, performing error-and-erasure correction decoding is a way to increase code performance gains without incurring increased cost in system complexity. The mechanism we have chosen for erasing unreliable NC-BFSK symbols is based on a ratio threshold test, in which channel symbols having a signal envelope ratio (i.e. the ratio between the decision variables) below a certain threshold are erased. This erasure declaration mechanism does not use any channel state information; hence, requires very little additional complexity in comparison to hard decision decoding. In the branch metric computations, erasures are

assigned a value half-way between the binary values for the expected symbols.

We also investigated error-and-erasure correction decoding for RS codes, which achieves some performance benefit with a trivial increase in decoding computation in comparison to error correction decoding [17]. The mechanism for declaring erasures is again based on a ratio threshold test, in which the code symbols having the lowest signal envelope ratio, z_0/z_1 (assuming that $z_0 > z_1$), is erased. When $z_0/z_1 > \theta > 1$, the decision corresponding to z_0 appears to have a good quality. This erasure declaration metric recognizes that the worst BFSK symbol in each q-bit code symbol is the weak link but it will inevitably fail to erase some symbols which are in error, and will erase some symbols which are not in error. There is an optimum range of values for the number of erasures, N_e, declared so that residual error correcting capability, N_c, is sufficient to correct the remaining errors in the received word. The optimum range of values was found during simulation.

4.5 Performance Evaluations

Performance evaluation by simulation for a NC-BFSK system over fading AWGN channels, as well as, multiple-access channels using soft decision decoding metrics given in (9), (10), (11), (14) and (15) was performed. Receivers with perfect knowledge of channel parameters α and N_o, and ones with estimated channel parameters were simulated. Table 2 summarizes the relevant E_b/N_o data for a rate-1/2, 32-state convolutional coded system with dual antenna diversity.

Table 2: Eb/No at Pb = 10^{-3} for different decoding metrics for a dual antenna diversity system.

	Fading AWGN Channel		Multiple-Access Channel	
Metric Type	Perfect	Estimated	Perfect	Estimated
Maximum Likelihood	8.25	9.25	8.25	9.10
Linear Combining	8.50	8.50	10.0	10.0
Euclidean	8.25	9.0	8.75	9.50

The data in Table 2 shows that there is no single metric that out-performs all others in every scenario. For the fading AWGN channel, the linear combining metric is clearly a very good selection in terms of having a good complexity-performance tradeoff. However, for the multiple-access channel, the maximum-likelihood (ML) metric might be the better choice since it holds a slight performance advantage over the other two metrics. In short, when fading interferers are the dominant impairment, the metrics using imperfect estimates of channel parameters performed better than the one using no channel parameters at all. For single antenna systems, the performance gap between the alternative metrics is even bigger.

For comparing soft decision decoding against hard decision decoding, the simulation showed that soft decoding with the ML metric performed 3.5 dB better than hard decision decoding for a rate-1/2, 32 state convolutional code for a single antenna sys-

tem. When dual antenna diversity was employed, the improvement decreased to approximately 2 dB. Fig. 3 illustrates the improvements in SNR of soft decision decoding over hard decision decoding for SNR ranging from 8 to 20 dB.

Simulation of the error-and-erasure correction decoding for a RS code was also carried out and the results are included in Table 3, where performance of coded systems for different coding schemes and decoding methods are summarized.

Table 3: Code Performance for Different Decoding Methods

Code Type RS = Reed-Solomon CC = Convolutional	Decoding method	Gain w/Single Antenna (dB)	Gain w/Dual Antenna (dB)
RS (n=31, k=15, d=17)	Error Correction	13.1	18.2
	Error & Erasure Correction	14.9	19.5
CC (r=1/2, v=5)	Hard Decision Decoding	13.5	19.3
	Error & Erasure Decoding	15.0	20.0
	Soft Decision Decoding	17.0	21.5

The simulation data showed that a rate-1/2, 32 state convolutional code with soft decision decoding requires approximately 2 dB less signal-to-noise at 10^{-3} BER than a length 31, RS code with error-and-erasure correction decoding on a Rayleigh fading channel, both with and without dual antenna diversity. Thus, a convolutional code with ML soft decision decoding appears to be the most suitable code selection for our particular frequency hopped wireless architecture.

5. Conclusion

We have considered channel coding options for a frequency-hopped system employing NC-BFSK signaling. We have concluded that the hopping patterns should be selected so as to randomize the interference encountered in successive hops, and have observed the advantages offered by the Latin squares construction for synchronous systems. Alternative channel codes were compared on the basis of delay, complexity, and bandwidth efficiency, with the conclusion that when hard decision decoding is employed, the BCH, RS, and convolutional codes that can meet the delay and bandwidth constraints have similar performance and comparable complexity. At the desired bit error rate, the convolutional codes are slightly better since they required approximately one dB less signal-to-noise ratio than the best performing block codes. Furthermore, the performance gain obtained by soft decoding of the convolutional codes with the maximum-likelihood metric was shown to be more significant than the gain obtained by error-and-erasure correction decoding of block codes. This was true even with imperfect channel state estimates. The benefits of dual antenna diversity in combination with coding were demonstrated.

One way to form an estimate of the SNR for a slot is to accumulate the mean

squared error between the received signal and hard decision demodulation. The larger the mse, the less reliable the slot. The selection of a soft decision metric depends on the reliability of the channel parameters estimator. That is, the number of levels of quantization to be used in subsequent decoding depends on the application. For example, with a very slowly changing channel, results could be accumulated over several slots, and many bits of soft decision information extracted. At the other extreme, for only a small number of data symbols per slot the best that can be expected is to be able to declare erasures. The same is true of fast hopped systems, where all that could be done is to monitor the received signal strength and declare erasures when it is below some empirically determined threshold. In the future, we hope to investigate algorithms for adapting the metric to changing channel conditions and applications.

References

[1] J. Min, A. Rofougaran, V. Lin, M. Jensen, A. Abidi, G. Pottie, and Y. Rahmat-Samii, "A Low-Power Handheld Frequency-Hopped Spread Spectrum Transceiver Hardware Architecture," Virginia Tech Symposium on Wireless Personal Communications, Blackburg, VA, June 1993.

[2] G. Zaharia, G. El Zein, J. Citerne, "Time delay measurements in the frequency domain for indoor radio propagation," IEEE Antennas and Propagation Society International Symposium. 1992 Digest., pp. 1388-91

[3] D.M.J. Devasirvatham, C. Banerjee, R.R. Murray, D.A. Rappaport, "Two-frequency radiowave propagation measurements in Brooklyn," 1st International Conference on Universal Personal Communications '92 Proceedings, pp. 01:05/1-5.

[4] C. Chien, P. Yang, E. Cohen, R. Jain, and H. Samueli, "A 12.7 Mchips/s all-digital BPSK direct sequence spread-spectrum IF transceiver in 1.2 μm CMOS," To appear in 1994 IEEE International Solid-State Circuits Conference.

[5] G. Chang, A. Rofougaran, M-K. Ku, A.A. Abidi, H. Samueli, "A low-power CMOS digitally synthesized 0-13 MHz sinewave generator," To appear in 1994 IEEE International Solid-State Circuits Conference.

[6] B. Gudmundson, J. Sköld, and J.K. Ugland, "A Comparison of CDMA and TDMA Systems," Proc. 1992 IEEE Vehic. Tech. Conference, Denver, May 11-13, 1992, pp. 732-735.

[7] H. Sasaoka, "Block coded 16-QAM/TDMA cellular radio system using cyclical slow frequency hopping," Proc. 1992 IEEE Vehic. Tech. Conference, pp. 405-408.

[8] N.R. Livneh, R. Meridan, M. Ritz, and G. Silbershatz, "Frequency hopping CDMA for digital radio," Proc. Int'l. Commsphere Symposium, Herzliya, Israel, Dec. 1991.

[9] G.J. Pottie and A.R. Calderbank, "Channel Coding Strategies for Cellular Radio," IEEE International Symposium on Information Theory, San Antonio, TX, Jan. 1993. (submitted to IEEE Transactions on Vehicular Technology).

[10] A.D. Wyner, "Multi-tone multiple-access for cellular systems,", AT&T Bell Laboratories Tech. Memo, BL011217-920811-11TM, Nov. 1991.

[11] J. Roche, "Families of binary matrices having nearly orthogonal rows," AT&T Bell Laboratories Tech. Memo, BL011217-920812-12TM

[12] K.S. Gilhousen, et. al., "On the Capacity of a Cellular CDMA System," IEEE Trans. Vehic. Tech., vol. 40, May 1991, pp. 303-312.

[13] J.G. Proakis, *Digital Communications*. 2nd Ed. New York: McGraw-Hill, 1989.

[14] A.J. Viterbi, J.K. Omura, *Principles of Digital Communication and Coding*. New York: McGraw-Hill, 1979.

[15] S. Lin and D.J. Costello, *Error Control Coding: Fundamentals and Applications*. Englewood Cliffs, NJ: Prentice-Hall, 1983.

[16] A.D. Whalen, *Detection of Signals in Noise*. San Diego: Academic Press, Inc., 1971.

[17] R.E. Blahut, *Theory and Practice of Error Control Codes*. New York: Addison-Wesley, 1983.

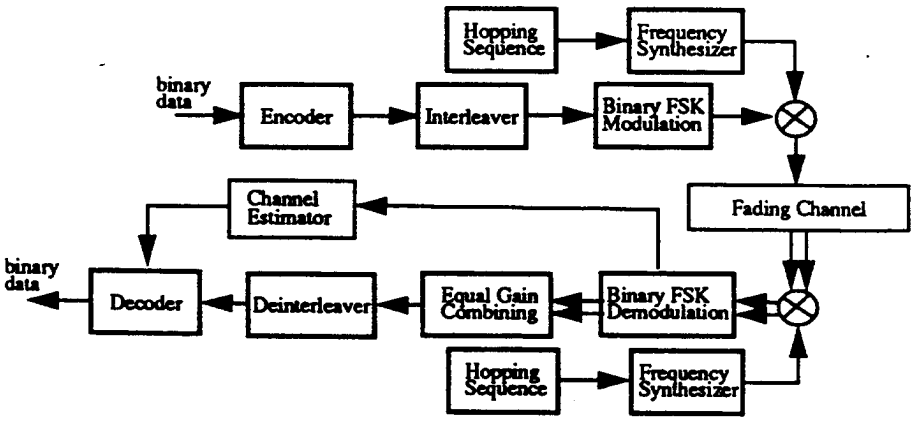


Fig. 1. System Model.

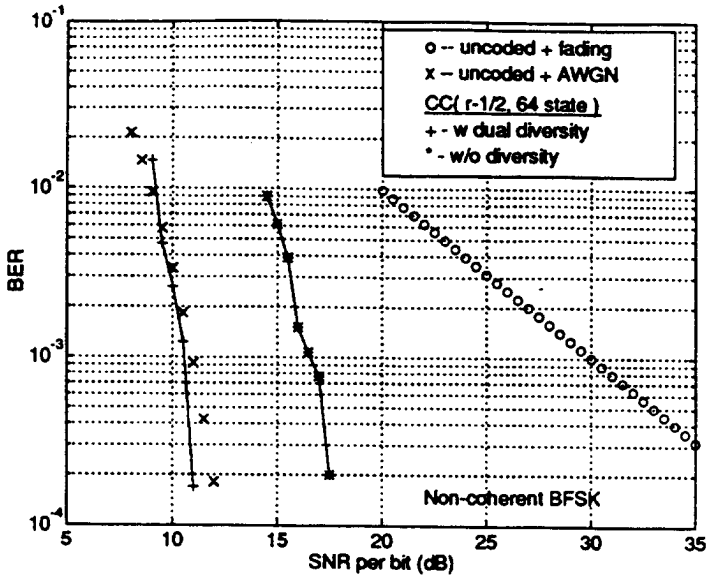


Fig. 2. BER vs. SNR for rate-1/2, 64 state code.

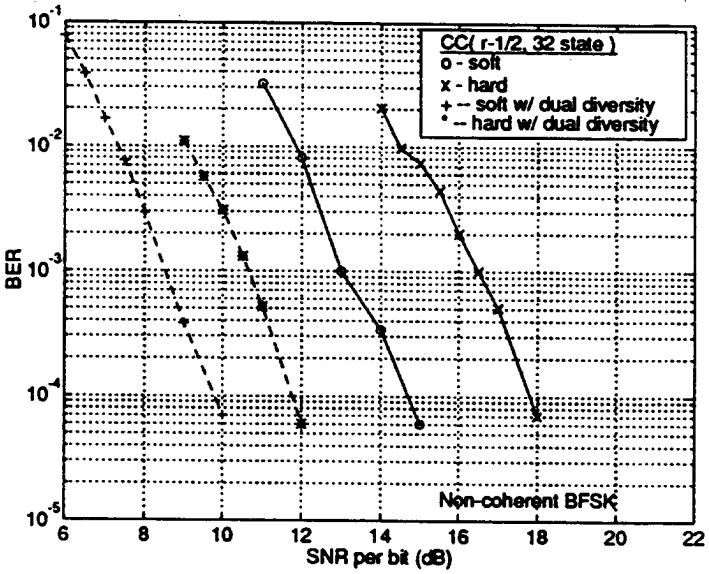


Fig. 3. BER vs. SNR for rate-1/2, 32 state code.

A Low-Power Handheld Frequency-Hopped Spread Spectrum Transceiver Hardware Architecture

J. Min, A. Rofougaran, V. Lin, M. Jensen, H. Samueli, A. Abidi,
G. Pottie, and Y. Rahmat-Samii

Wireless Personal Communications, Trends, and Challenges
chapter 11, pp. 119-126
Kluwer Academic Publishers, 1994

11

A Low-Power Handheld Frequency-Hopped Spread Spectrum Transceiver Hardware Architecture

Jonathan Min, Ahmadreza Rofougaran, Victor Lin, Michael Jensen, Henry Samueli, Asad Abidi, Gregory Pottie, and Yahya Rahmat-Samii

Integrated Circuits and Systems Laboratory
Electrical Engineering Department
University of California, Los Angeles

Abstract - The overall goal of this research project is to develop low-power personal communications transceiver hardware technologies, coupled with advanced systems techniques such as antenna diversity, channel coding, and adaptive power control, for achieving robust wireless digital data transmission over multipath fading channels. A frequency-hopped spread spectrum (FH/SS) code division multiple access (CDMA) technique was chosen over other multiple access schemes because a) it provides an inherent immunity to multipath fading; and b) the signal processing is performed at the hopping rate, which is much slower than the chip rate encountered either in a direct sequence (DS) CDMA or time division multiple access (TDMA) system, thereby potentially resulting in much lower receiver power consumption. Furthermore, frequency-shift keying (FSK) modulation with noncoherent detection in a frequency-hopped system results in a much simpler transceiver architecture as compared with coherent amplitude and phase modulation methods commonly used in DS and TDMA systems. Architectural innovations as well as advanced circuit techniques will be incorporated into the design of a portable handset to achieve minimum power and size without sacrificing robustness in performance. A 3-V CMOS implementation is being developed for the entire transceiver including the 900 MHz radio frequency (RF) front-end. Multiple miniature antennas will also be integrated into the handset to achieve the maximum diversity benefit for robust data transmission. The hardware and system technologies developed for the transceiver design will be general enough to be applicable to a wide variety of commercial and military wireless communications applications such as cellular and micro-cellular radios and telephones, wireless LANs, and wireless PBX systems. In order to validate the design techniques being proposed, a prototype all-CMOS transceiver handset with dual antenna diversity is being developed to demonstrate pedestrian-to-pedestrian communication over the 902-928 MHz band. This paper overviews the proposed FH/SS transceiver hardware architecture. The system rationale, analog and digital circuits, and antenna design techniques for the transceiver are also presented in the following sections.

I. System Architecture

1) System Rationale

The UCLA personal communications project has the objective of passing data at rates up to 160 kb/s between two low-power handsets, using the 902-928 MHz band. The principal channel impairments in multi-user radio systems are attenuation due to shadowing, multipath fading, and interference from other radios. To overcome these impairments without resorting to high transmitter power, the proposed architecture incorporates many advanced techniques. Two antennas will be used to provide polarization/space diversity. Frequency hopping combined with error control coding will be implemented to provide both frequency and interferer diversity, further reducing the effects of multipath fading. Adaptive power control will be employed so that the minimum transmitted power required for reliable communication is used. Finally, an adaptive data rate permits a longer transmission range for lower rate messages, without increas-

ing the peak transmit power.

The most damaging type of fading is Rayleigh fading, and it is also frequently encountered. Whereas in a Gaussian channel the bit error rate (BER) for a well-designed system drops exponentially with SNR, for a Rayleigh fading channel the BER declines only linearly with SNR [1]. Thus, a huge power penalty must be paid unless diversity techniques are used to mitigate the effects of multipath propagation. Diversity is the technical term for reception of different versions of the same information, with different fading levels. With L^{th} order diversity, the probability of error declines as the L^{th} power of SNR. Diversity may be achieved in any combination of the space (antenna), time, or frequency domains. A combination of techniques for each domain provides robust performance and is an economical means of achieving a high aggregate diversity order. This permits receiver performance close to that achievable in the additive white Gaussian noise channel.

Frequency diversity may be achieved by hopping many times per symbol. In a synchronous system, it is possible to arrange the hopping patterns so that no interference is generated by users within the same cell, and in addition, the hopping patterns will intersect those of users in surrounding cells in only one position. Thus, there will be no dominant interferer. At higher data rates, it may only be possible to hop but once per symbol. In this case, channel coding is essential in providing frequency and interferer diversity protection [2]. Furthermore, channel coding also increases the effective diversity order for faster hopped systems (multiple hops per symbol), as well as providing the usual benefits of coding gain.

Frequency-hopped systems permit a large frequency/interferer diversity to be obtained with lower transmission delay and significantly less complexity than wideband direct sequence spread spectrum or time division multiple access systems. However, fast hopping generally precludes the use of a coherent receiver. Consequently, the preferred modulation is binary FSK. This implies a significant SNR versus BER performance penalty--roughly 6 dB as compared to coherent binary PSK. However, the combination of frequency diversity, synchronous access and reduced complexity compared to DS/SS CDMA more than makes up for this penalty.

2) Channel Modeling and System Simulation

The radio propagation channel can be conceptualized as a time-variant linear filter that transforms input signals into time-varying output signals. Various research groups have proposed radio channel models for use in simulations [3]. In our initial investigation, the tapped-delay line model methodology was chosen and a Rayleigh fading channel was implemented. A typical application for a propagation model is to estimate the coverage area of the communication system. In our preliminary link budget analyses several scenarios were considered. The results for a portable-to-portable communication link are summarized in Table 1.

Table 1. Estimated Coverage Areas for a 900 MHz Radio Link Between Two Portable Units

OPERATING CONDITION	RADIUS OF COVERAGE
Clear Line-of-Sight (LOS) Channel	1470 ft
LOS Channel with Tree Obstructions	620 ft
Non-LOS Channel without Tree Obstructions	695 ft

In the link budget calculations, it was assumed that the transceiver employed two branches of antenna diversity to mitigate multipath effects. Other assumptions included FSK modulation, an average BER of 10^{-3} to provide good-quality voice reproduction, and an average transmit power of 10 mW. A portable to base-station communication scenario was also considered as a way of improving the coverage area due to the improved base-station antenna performance. This situation is similar to that of the mobile cellu-

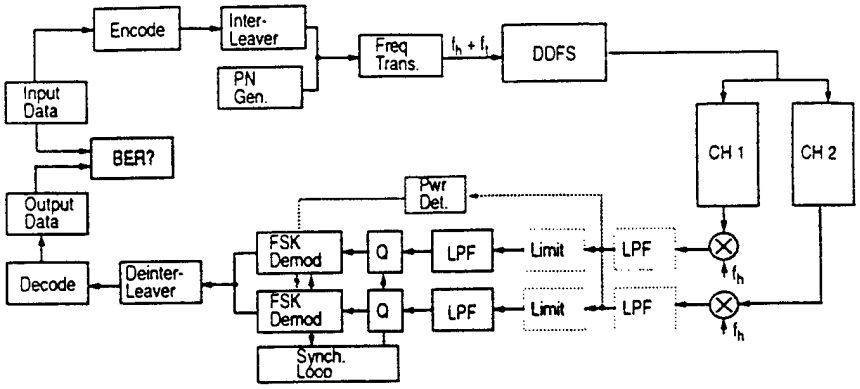


Figure 1. Simulation block diagram.

lar environment. It was found that a substantial improvement can be achieved by using a base-station to act as a switching office for relaying messages between portables. For example, by utilizing a base-station in the non-LOS channel scenario considered above, the radius of coverage increases to around 1500 feet.

The overall system simulation including the radio fading channel is in progress to evaluate the performance of the frequency-hopped spread spectrum (FH/SS) transceiver. System features such as antenna diversity and channel coding will be incorporated into the transceiver architecture and synchronization simulations. Various diversity combining techniques will be compared to trade the performance gain versus the hardware complexity. Simulations will be performed over a 26 MHz hopping bandwidth, using VANDA (Visual Analysis and Design Automation) [4], an integrated CAD environment for communication signal processing systems being developed at UCLA. The overall simulation block diagram is shown in Figure 1.

II. Hardware Architecture

1) RF/IF Architecture

A baseline RF/IF architecture employing high-performance low-power circuit innovations has been proposed for implementing the FH/SS transceiver. The combination of a direct digital frequency synthesizer (DDFS) and a digital-to-analog converter (DAC) is used to implement the fast frequency hopping. In order to simplify the designs of the DAC, the image reject filtering, the local oscillator, and the up conversion circuits, the DDFS is designed to hop only in a 0-13 MHz band. A 26 MHz hopping bandwidth is obtained by frequency doubling in the following RF stages. In addition, because of the doubling, the frequency of local oscillator (LO) used for up conversion will be 451 MHz, not 902 MHz, which results in power savings.

When the signal is up-converted, the SSB modulation technique [5] is used to reject the image of the signal. Image rejection can be achieved by taking either the real or imaginary part of the multiplication of two complex signals. This eliminates the need for filters with sharp roll-offs, and therefore lowers the overall power consumption. Image reject filtering is easily achieved in an FSK system because both the in

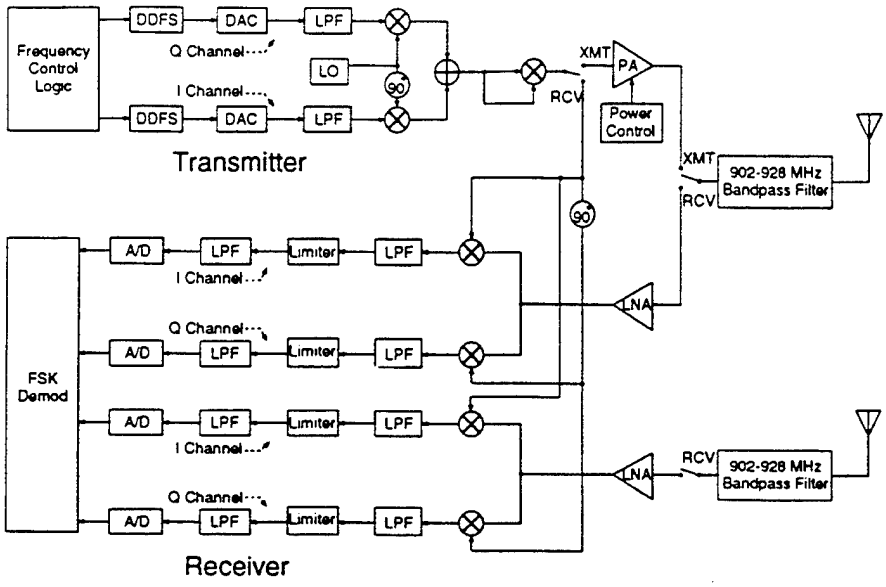


Figure 2. FH/SS RF/IF block diagram.

phase and quadrature signal components of the transmitted signal are easily and precisely generated using a quadrature DDFS, as shown in Figure 2. After rejecting the images, frequency doubling is used to convert the signal to the desired RF band (900MHz). Since we are using an FSK modulation scheme, a simple highly efficient class-C power amplifier can be used for transmission. To further minimize power consumption, the power amplifier will be integrated into a single chip along with the rest of the transceiver circuits. Finally, a low-cost off-the-shelf 902-928 MHz dielectric resonator bandpass filter will be used between the power amplifier and the antenna to reject the out-of-band harmonics and meet FCC transmission mask requirements.

Before down conversion, the received signal needs to be amplified using a low noise amplifier (LNA). The gain and noise figure of the LNA are critical factors in determining the overall transceiver performance. We have successfully developed a new technique to achieve a monolithic high-Q inductor in CMOS without altering the fabrication process [6]. This technique will be applied to design an LNA with high gain at 900 MHz. The high-Q inductor amplifier also provides extra filtering, thereby further attenuating out-of-band signals.

To minimize the number of high frequency components and thereby save power, the de-hopping and down conversion is performed in a single "direct-conversion" stage. In other words, the 900 MHz signal is de-hopped and down-converted directly to baseband. The down conversion circuit must achieve a low intermodulation distortion (IMD). One way to reduce power and lower the IMD products is to use a sub-harmonic sampling circuit which is switched at an integer (small) fraction of the RF frequency (900 MHz). This method can easily be implemented using conventional switch-capacitor techniques which are well developed for CMOS technology. These techniques not only result in low IMD products but also result in lower overall power dissipation.

2) Baseband Architecture/Algorithms

One of the key requirements for the portable communicator is low power consumption. This requires that complex signal processing, such as adaptive equalization, be avoided if possible. The hopping rate will be carefully chosen to meet this requirement [7]. Quadrature demodulation is required to efficiently use the entire 26 MHz transmission bandwidth. The automatic gain control (AGC) function is implemented using a combination of a lowpass filter (LPF), a hard limiter, and another LPF. Due to odd harmonics produced by hard-limiting the sine wave, this approach limits the modulation scheme to binary FSK rather than M-ary FSK. However, the merits from power saving and hardware simplicity obtained from the absence of an expensive (in both power and complexity) linear variable gain amplifier (VGA) and relaxation of the analog-to-digital converter (ADC) resolution (4 bits only) justifies this architecture for a portable handset design. The quadrature binary FSK modulation scheme also relieves the problem of group delay variations at different frequencies in the anti-aliasing filter. Following the ADC, the baseband signal processing is all performed digitally. This includes binary FSK demodulation, frame synchronization, PN acquisition, clock recovery, frequency tracking, and forward error correction.

Several architectural trade-off studies have been performed to determine the baseband (BB) architecture for the transceiver. An all-digital quadrature demodulation architecture using a correlation detector has been chosen over a fast fourier transform (FFT) or matched filter method. This choice is based on the fact that the correlator provides a flexible design which can easily accommodate programmable data rates, and the same hardware can be used to demodulate both data and sync hops.

A simple frame structure for the time duplexing mode (half-duplex) has been developed for master/slave configured communication transceivers. The frame structure consists of three fields: C0, C1, and C2. C0 is for frame synchronization, C1 for receiver ID, and C3 for user data. Synchronization is accomplished in three steps. The slave handset first listens for the C0 field of the frame, which is broadcasted by the master. Once frame synchronization is achieved, PN code acquisition follows simply by detecting the C1 field to match the pre-assigned receiver ID pattern. The PN code is synchronous with the frame and the pattern restarts at the same point with every new frame. The time track function is achieved by detecting the timing error and correcting the numerically controlled oscillator (NCO) accordingly. The overall frequency error is assumed to be small and therefore frequency tracking may not be necessary in a noncoherent FSK system, however, a frequency tracking architecture will be developed as a backup.

III. Advanced Analog and Digital Circuits

Various low-power analog and digital communication transceiver building blocks are under development. They range from RF analog to baseband digital circuits, all developed using CMOS technology. A 3-V, 800-MHz tuned RF amplifier in 2-μm CMOS was designed, fabricated and characterized. Several other circuits which are currently in design or fabrication are described in this section.

RF Amplifier - A low-power, low-noise RF amplifier in 2-μm N-well CMOS has been implemented and tested [6]. The amplifier has been measured to provide 14 dB of gain while dissipating only 6.9 mW (excluding output buffers) with a single 3-V supply. The amplifier has been designed to be a tuned amplifier in order to limit the noise contribution from the amplifier itself. The tuning of the amplifier was achieved by a large on-chip spiral inductor resonating with the parasitic capacitances. A novel technique has been applied to suspend spiral inductors over air-filled pits in order to attain over 100 nH of inductance on chip while providing a high self-resonant frequency. The noise figure of this amplifier was measured to be 6.1 dB (excluding input terminations), and the center frequency was 770 MHz. A new version of this amplifier is being developed in 1-μm CMOS which will be a tuned amplifier centered at 915 MHz. A lower noise figure is expected, and special attention will be given to the matching circuits between the amplifier

inputs and the antenna.

Sub-Sampler Circuit - A low-distortion sub-sampler circuit using 1-μm CMOS technology is in fabrication. The sub-sampler is to be used in the receiver after the RF pre-amplifier stage to perform down conversion. To achieve the desired linearity performance, the design requires a high-speed op-amp and careful optimization of switch sizes and capacitance values. Simulation results from extracted layout show that the op-amp can achieve a 70 dB DC gain and a 500 MHz unity-gain frequency when driving a 300 fF load (with a 60° phase margin). Moreover, the power dissipation can be kept under 10 mW with a single 3-V power supply. Sub-sampler circuit simulations also show 11-bit linearity when sampling an AM signal at 40 MHz (with a carrier frequency of 400 MHz).

Frequency Synthesizer (PLL/VCO) - A low phase noise clock source is required to up-convert and down-convert the signals in the transceiver. A frequency synthesizer (FS) must be designed to generate these high clock frequencies by scaling a stable reference source. The FS must not only have low phase noise but it must also have a low power dissipation. These specifications are further complicated by the large operating temperature range and low-power supply (3-V). A phase-locked loop (PLL) was chosen to keep the phase noise to a minimum. The individual circuit blocks will be optimized for low power dissipation. The PLL consists of an external crystal used to generate the reference frequency, a phase/frequency detector, a charge pump, a voltage controlled oscillator (VCO) and a digital divider. A low phase noise design should incorporate a high gain phase/frequency detector and a low gain VCO.

Digital-to-Analog Converter - A monolithic low-power 30 Msample/s 10-bit digital-to-analog converter (DAC) with a 3-V supply is being developed. Conventionally, high-speed and high-precision DACs are implemented with current mode architectures which implies that they are usually power hungry. The major challenge for this circuit lies in finding an architecture that will provide high speed while still maintaining high linearity. The proposed DAC employs a pipelined charge redistribution algorithm implemented differentially, in which the pipelined digital code controls MOS switches that charge the capacitors to some known reference voltage. Due to the inherent nature of a switched-capacitor circuit, there is no DC standing current running through the circuit itself. Thus, it is ideally suited for the low-power application we are seeking for this project. The analog signal will traverse a series of charge redistribution stages. After 10 such stages (10 bits), the final analog output is available, and it feeds directly to a sample and hold buffer that includes a low-power op-amp. In fact, the output buffer is the only active component of the DAC; the rest consists of only passive components such as MOS switches and capacitors. Due to the high speed and low supply voltage requirements, the analog output swing is limited to only 0.5 V. This small output swing will not degrade the performance since the total thermal noise of the capacitors is less than one-half an LSB.

Direct Digital Frequency Synthesizer - A low-power CMOS quadrature direct digital frequency synthesizer (QDDFS) has been designed and is currently in fabrication. The QDDFS synthesizes 10-bit output sine and cosine waves simultaneously at 40 Msample/s. The synthesizer covers a bandwidth from DC to 20 MHz with a switching speed of 25 ns and a tuning latency of 2 clock cycles. Several techniques are employed to reduce the ROM storage [8]. A general-purpose ROM contents generation/simulation program has been developed, which can be used to optimize the ROM contents to get the best spurious response and ROM size compression. The worst-case spurious response for the proposed DDFS is -72.63 dB, which is analytically guaranteed for all output frequencies.

IV. Miniature Antenna Design

The design of miniature antennas suitable for implementation on a handheld transceiver involves investigation of radiating elements which can achieve maximum diversity performance when confined to a unit whose dimensions are less than a wavelength. This investigative process requires a detailed theoretical

examination and comparison of the behavior of a wide variety of candidate antenna geometries operating in the presence of the transceiver structure. To accomplish this task, simulation tools based upon the finite-difference time-domain (FDTD) algorithm [9] and moment method have been developed which provide high accuracy while maintaining the flexibility required to model different antenna configurations.

These state-of-the-art antenna simulation tools have been used to investigate the performance of several different antenna topologies mounted on a conductive transceiver housing. The formulation allows determination of the antenna radiation pattern, gain, input impedance, and envelope correlation coefficient (diversity performance) over a wide frequency band. Figure 3 shows a representative example of the results of the FDTD analysis tool for the case of a Planar Inverted-F Antenna (PIFA)--an air-substrate modified microstrip antenna which provides efficient radiation for small physical dimensions [10]. This figure illustrates the antenna/transceiver geometry with its associated radiation pattern (normalized to the directivity) and input impedance. Numerous similar computations are being performed for other antenna geometries in an effort to identify an optimal diversity configuration.

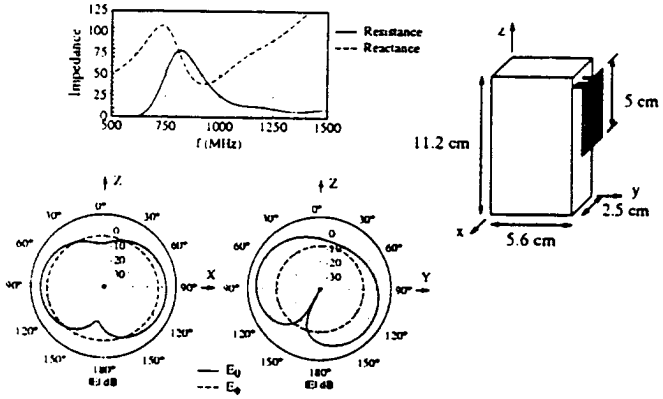


Figure 3. Planar Inverted-F Antenna on a conducting transceiver housing: geometry, input impedance versus frequency, and directivity pattern.

V. Conclusion

In this paper, the overview of the UCLA FH/SS transceiver project has been presented, from system issues to miniature antenna design. Robust operation of the handset will be achieved through the use of space diversity (dual antenna), frequency diversity (spread spectrum), and time diversity (channel coding). An efficient hardware architecture using FH/SS techniques has been proposed, utilizing direct conversion, frequency doubling, sub-harmonic sampling, and parallel interleaving. An open-loop hard-limiting automatic gain control (AGC) method is chosen to simplify hardware and reduce power consumption. Various RF/BB building blocks for the transceiver have been identified and are being developed. Highly integrated design of these low voltage custom analog and digital CMOS components reduces the use of discrete components, which enables the transceiver to achieve the minimum power consumption.

Acknowledgment

This work is supported by DARPA contract DAAB07-92-R-C977.

References

[1] J.G. Proakis, *Digital Communications*, 2nd Ed. New York: McGraw-Hill, 1989, pp. 719-728.

[2] G.J. Pottie and A.R. Calderbank, "Channel Coding Strategies for Cellular Radio," *IEEE International Symposium on Information Theory*, San Antonio, TX, Jan. 1993.

[3] T.S. Rappaport, "Statistical Channel Impulse Response Models for Factory and Open Plan Building Radio Communication System Design," *IEEE Transactions on Communications*, pp. 794-807, May 1991

[4] P. Tjahjadi, P. Yang, B. Wong, B. Chung, E. Cohen, and R. Jain, "VANDA - A CAD System for Communication Signal Processing Circuits Design," *VLSI Signal Processing IV*, pp. 43-52, IEEE Press, 1990.

[5] P. O'Sullivan, R. Benton, A. Podell, and J. Wachsman, "Highly Integrated I-Q Converters for 900MHz Applications," in *Technical Digest of IEEE 1990 GaAs IC Symposium*, New Orleans, LA, pp. 287-290.

[6] J.Y-C. Chang and A.A. Abidi, and M. Gaitan, "Large Suspended Inductors on Silicon and Their Use in a 2-μm CMOS RF Amplifier," to appear in *IEEE Electron Device Letters*, vol. 14, no. 5, May 1993.

[7] A.A.M. Saleh, A.J. Rustako, L.J. Cimini, G.J. Owens, and R.S. Roman, "An Experimental TDMA Indoor Radio Communications System Using Slow Frequency Hopping and Coding," *IEEE Transactions on Communications*, vol. 39, pp. 152-162, Jan. 1991.

[8] H.T. Nicholas and H. Samueli, "A 150-MHz Direct Digital Frequency Synthesizer in 1.25-μm CMOS with -90 dBc Spurious Performance," *IEEE Journal of Solid-State Circuits*, vol. 26, pp. 1959-1969, Dec. 1991.

[9] M.A. Jensen and Y. Rahmat-Samii, "FDTD Analysis of PIFA Diversity Antennas on a Handheld Transceiver Unit," *IEEE Antennas and Propagation Society International Symposium*, Ann Arbor, MI, Jun. 1993.

[10] T. Taga and K. Tsunekawa, "Performance Analysis of a Built-in Planar Inverted-F Antenna for 800MHz Band Portable Radio Units," *IEEE Journal on Selected Areas in Communications.*, vol. SAC-5, pp. 921-929, 1987.

A Highly Linear 1-GHz CMOS Downconversion Mixer

P.Y. Chan, A. Rofougaran, K.A. Ahmed, and A.A. Abidi

19th European Solid-State Circuits Conference Proceedings
pp. 210-213
September 1993.

A Highly Linear 1-GHz CMOS Downconversion Mixer

P. Y. Chan, A. Rofougaran, K. A. Ahmed, and A. A. Abidi

Integrated Circuits & Systems Laboratory
University of California
Los Angeles, CA 90024-1594

ABSTRACT

A highly linear, closed loop CMOS architecture is demonstrated for the mixing of RF waveforms to baseband as sampled-data signals. In converting a 915 MHz RF waveform to a 20 MHz spread spectrum baseband signal, the measured third-order intercept lies at +27 dBm of input power. As the 1-μm CMOS prototype dissipates only 12 mW from a 5V supply, and is capable of operation at 3V, it is expected to be of use in advanced handheld wireless receivers.

Introduction

IC design faces a new challenge posed by the emerging requirements of advanced transceivers for wireless communications. First radio pagers, and then cellular telephones, have prompted highly integrated solutions for the required RF and IF components. It is widely felt today that ICs in advanced bipolar and GaAs technologies will provide the best solutions in terms of the sought performance at the lowest power dissipation. Several chipsets have been developed in these technologies [1-3]. We believe that for operation at RF in the 1 GHz band, even CMOS is a viable alternative. A highly integrated CMOS RF and IF chip eliminates the need to route signals on low impedance lines between chips, thus saving the power wasted in buffers. Our circuit designs are based on a total system solution for a frequency-hopped spread-spectrum transceiver, which optimally partitions signal processing between the analog and digital domains.

For use in the receiver, we have already demonstrated a low power 750 MHz CMOS RF amplifier with on-chip tuning elements [4]. The first mixer in the receiver poses an even a greater design challenge, because intermodulation distortion between an interference signal passing through the preselection filter and the desired signal may produce in-band spurious signals, and irrecoverably degrade the received signal to noise ratio. Thus, a mixer with a very wide dynamic range is sought. Mixer design is today a well acknowledged art unto itself.

Mixers: Bipolar, Switching, and CMOS

Conventional mixers for continuous-time waveforms downconvert by producing an output at the difference frequency between the RF and local oscillator (LO) inputs. For direct conversion to baseband, the LO is close in frequency to the RF, and thus itself at a high frequency. A mixer is more linear, and more effectively suppresses intermodulation distortion, either by how closely its input-output relationship approaches an analog multiplication, or if it is a commutating switch, by the strength of LO driving the switch and its turnoff aperture time [5]. For these reasons, commutating mixers do not often find use in low power, battery operated receivers, although they are quite common as standalone microwave components.

Mixers based on Gilbert's analog multiplier are ubiquitous in bipolar ICs [1,5]. Intermodulation distortion arises in this circuit because at RF it is difficult to predistort the input signal to cancel the exponential bipolar transistor characteristic, and also from imperfections such as the large base spreading resistance found in transistors with a high cutoff frequency. Nonlinearity in a balanced mixer is specified by the third-order intercept (IP3), the RF input power at which the third harmonic strength at the output is equal to the fundamental frequency. The larger the IP3, the lower the intermodulation distortion. The IP3 of mixers based on the Gilbert multiplier is usually somewhere below 0 dBm, and if simple circuits compensate some of the imperfections, it may be raised to 5 or perhaps 10 dBm. This is only marginally acceptable in spread spectrum receivers, which may be susceptible to very large narrowband interfering signals.

When the Gilbert cell topology is used in a GaAs FET mixer, cross-modulation terms appearing in the circuit input-output characteristic result in a lower IP3 than in its bipolar antecedent. It is not uncommon in GaAs ICs to mix signals using the theoretical quadratic I-V characteristic of a single FET. However, owing to short channel effects and other physical phenomena, the non-ideal characteristics of physical FETs will produce additional, often strong, cross-modulation terms, which will limit mixer linearity. A CMOS mixer with this topology, if operational at all at 1 GHz, will be even more nonlinear than the GaAs FET circuits.

An altogether different tack was taken in this design by letting the mixer produce a sampled data, discrete-time analog output signal. This may then be encoded by an A/D converter for use in a digital receiver. The design stems from the realization that the track and hold (T/H) circuit block is itself a mixer: it mixes the input signal with the sampling clock. A T/H clock frequency equal to the baseband sample rate thus directly samples the RF to produce a discrete-time analog downconverted signal at the output (Fig.1). Extracting narrowband information about the modulation by subsampling a high frequency carrier is finding use in instrumentation [6], and even in communication circuits [7].

Closed-loop T/H circuits using op amps are commonly used in CMOS to attain high linearity and cancel charge feedthrough from MOS switches (Fig.2). A properly designed T/H presents a simple *open-loop RC* circuit to the input in *track* mode, and applies feedback to the sampling capacitor only in hold mode. The track mode bandwidth is limited by the ON resistance of FET switches, FET capacitance, and the sampling capacitor. This will be well above 1 GHz in a 1-μm CMOS technology. After faithfully tracking an RF signal, the sampling capacitor acquires its instantaneous value when the sampling switch turns OFF. The acquisition aperture of the FET switch determines the effective sampling bandwidth. In the limiting case when an infinitely fast clock edge turns it OFF, the aperture is the time it takes to empty an inverted FET channel of carriers. This may be as fast as a few tens of picoseconds in a FET of 1-μm channel length. When the transition from track to hold mode is made with the proper clock phases, the charge feedthrough from the switches on to the sampling capacitor is either entirely removed, or appears as an input common mode jump, without causing any distortion on the held sample in either case [8].

There are several advantages in using a subsampling mixer over a conventional mixer in a wireless receiver. As the LO is at the baseband sampling frequency, and not close in frequency to the RF, there is almost no spurious radiation caused by LO leakage through the antenna. Furthermore, this subsampling mixer requires a large LO drive on the switches for linear operation, but unlike a conventional commutating mixer where a high frequency amplifier must boost the LO level, here a rail-to-rail voltage swing is readily obtained at the baseband LO frequency.

Circuit Design

The mixer is a closed-loop track and hold which operates in two frequency ranges widely spaced apart between these two modes. The op amp must settle fully within 10 ns to support a 50 MHz sample rate, and for low distortion, its DC gain must be about 60 dB. A super-cascode topology (Fig.3) best satisfies these requirements [9]. NMOS common source auxiliary amplifiers boost the DC gain of the cascode FETs. While meant to operate nominally at 5 V, the amplifier is fully functional at 3 V. FET sizing is optimized for the shortest settling time when driving the specified capacitances at the input,

output, and in feedback. SC common-mode (CM) feedback applied to the tail current source forces the bias value of the output nodes to a reference CM voltage.

In track mode, the differential input and output terminals of the op amp are shorted by switches, thus disconnecting it entirely from the signal path (Fig.2). The desired input CM reference level is applied to the input terminals. Thus in track mode, the RF signal encounters a circuit consisting only of three switches and two 500 fF capacitors. The switch FETs are sized to obtain a low ON resistance (75 Ω) when driven by full logic levels at their gates, enabling the capacitor to track the input voltage. At the transition from track to hold, NMOSFET S1 and PMOSFETs S2 turn OFF first, followed by S3. Charge feedthrough from S1 causes a signal independent jump at the op amp CM input, which may drive the input terminals outside the CM input range, leading to a long recovery time. This problem is alleviated by sizing switches S2 so that their signal independent negative charge feedthrough approximately cancels the charge from S1. Uncancelled charge will however not cause any distortion. The *signal dependent* feedthrough on to the sampling capacitors from S3 when it turns OFF is removed by the negative feedback once the op amp output has settled. Thus, an accurate sample of the RF waveform appears at the op amp output.

The switches are driven by clock buffers consisting of a cascade of two CMOS inverters. Simulated falltime of the clock signals is 500 ps nominal, 800 ps worst-case over process and temperature, which is adequate for the sampling aperture sought. The circuit requires only two clock phases, one of which is delayed to produce the third clock. The clocks applied to sampling switches S1 and S2 must be well balanced.

Although the core of the mixer consumes a very small amount of power, its output is incapable of driving the 50Ω impedance of the measuring instruments. Therefore, an integral part of the design is an on-chip buffer to drive this impedance which is at least as linear as the mixer. A balanced differential pair buffer driving off-chip at the drains, degenerated by a polysilicon resistor at the sources, and enclosed within the op amp feedback was used (Fig.4). This unity gain buffer also converts the balanced mixer output to a single-ended signal to drive the instrumentation. While the mixer dissipates only 12 mW, the buffer dissipates 100 mW and it may swing as large as 1 V differential signal into 50Ω. On-chip 50Ω polysilicon resistors are also used to terminate the differential input pads.

Experimental Results and Discussion

The mixer occupies an active area of 0.4 × 0.65 mm on the die, and is fabricated in a double metal, single poly CMOS process with 1-μm gate feature size (Fig.5). The sampling capacitance is implemented between the two layers of metal. All measurements reported below are made with a 5 V power supply.

Using a mixer clock frequency (LO) which is a non-integer divisor of the input RF, a beat frequency is produced at the mixer output. Measurements are made at LO frequencies up to 120 MHz, and a variable output beat frequency as large as 50 MHz. Distortion in the discrete-time resetting buffered output is measured on a spectrum analyzer. All the key harmonics lie in the buffer passband. The measured IP3 at a 50 MHz sample rate is about +27 dBm with a 900 MHz RF input (Fig.6), and fundamental component at the output is compressed by 1 dB at an input power greater than +12dBm. These are some of the best results to our knowledge for an IC mixer in any technology operating in this RF band [1-3, 10], a comparable performance only available in discrete passive mixers using a transformer coupled diode bridge. The measured mixer conversion loss is 6 dB. As the mixer samples wideband noise across a 1 GHz input bandwidth, but then aliases that noise into a relatively narrowband output (determined by half the sample rate), the noise figure is 18 dB, very close to prediction. In the intended application, the mixer will be preceded by an RF amplifier with a 14 dB gain and 6 dB noise figure.

We have described a new use of a CMOS SC track and hold circuit as a downconversion mixer in a radio receiver. As the mixer plays a key role in determining the interference rejection of spread spectrum receivers, we foresee new possibilities arising from these striking results to fabricate receivers

operating in certain frequency ranges, such as in the unlicensed 902-928 MHz band, entirely as single-chip CMOS ICs.

Acknowledgement: This research was supported by ARPA, Rockwell International, and the State of California MICRO Program.

1. Wilson, J.F., *et al.*, "A Single-Chip VHF and UHF Receiver for Radio Paging", *IEEE J. of Solid State Circuits*, v. 26, pp. 1944-1950, December 1991.
2. Tokumitsu, M. and M. Muraguchi, "A Very Low Power Dissipation Front-End MMIC for L-Band Receivers", in *22nd European Microwave Conference*, Espoo, Finland, 1992.
3. Imai, Y., M. Tokumitsu, and A. Minakawa, "Design and Performance of Low-Current GaAs MMICs for L-Band Front-End Applications", *IEEE Trans. on Microwave Theory and Techniques*, v. 39, pp. 209-215, February 1991.
4. Chang, J. Y-C., A. A. Abidi, and M. Gaitan, "Large Suspended Inductors on Silicon and their use in a 2-μm CMOS RF Amplifier", *IEEE Electron Device Letters*, v.14, pp. 246-248, May 1993.
5. Chadwick, P.E., "High Performance Integrated Circuit Mixers", in *IEE Conf. on Mobile Radio Systems and Techniques*, pp. 1-9, London, 1984.
6. Ballo, D.J. and J.A. Wendler, The Microwave Transition Analyzer: A New Instrument Architecture for Component and Signal Analysis, in *Hewlett-Packard Journal*, October 1992, pp. 48-62.
7. Weisskopf, P.A., Subharmonic Sampling of Microwave Signal Processing Requirements, in *Microwave Journal*, May 1992, pp. 239-247.
8. Lin, Y.-M., B. Kim, and P.R. Gray, "A 13b, 2.5 MHz Self-Calibrated Pipelined A/D Converter in 3-μm CMOS", *IEEE J. of Solid State Circuits*, v. 26, pp. 628-636, April 1991.
9. Bult, K. and G. Geelen, "A Fast-Settling CMOS Op Amp for SC Circuits with 90 dB DC Gain", *IEEE J. of Solid State Circuits*, v. SC-25, pp. 1379-1394, December 1990.
10. Takenaka, T., S. Hara, and T. Tokumitsu, "A Miniaturized Broadband MMIC Mixer", in *GaAs IC Symposium*, pp. 193-196, 1989.

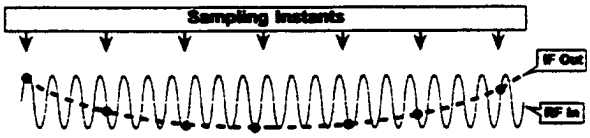


Fig.1. Samples of the RF input taken at an offset subharmonic frequency produce a discrete-time IF output.

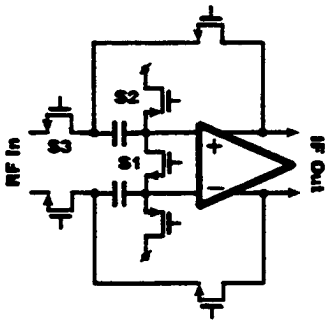


Fig.2. The downconversion mixer is a SC track and hold circuit with a very large tracking bandwidth.

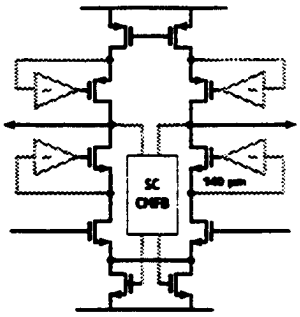


Fig.3. Super cascode op amp circuit.

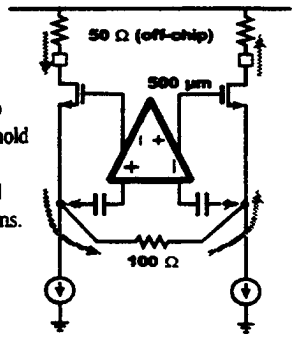


Fig.4. Mixer is embedded within a linear buffer circuit to drive off-chip instrumentation. Circuit shown in hold mode. With the on-chip input terminations, this becomes a useful standalone mixer for 50 ohm systems.

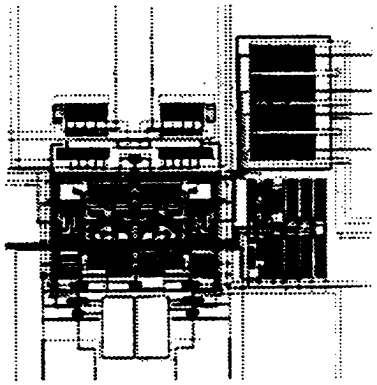


Fig.5. Microphotograph of active area of mixer. 500 fF double-metal sampling capacitors evident on bottom.

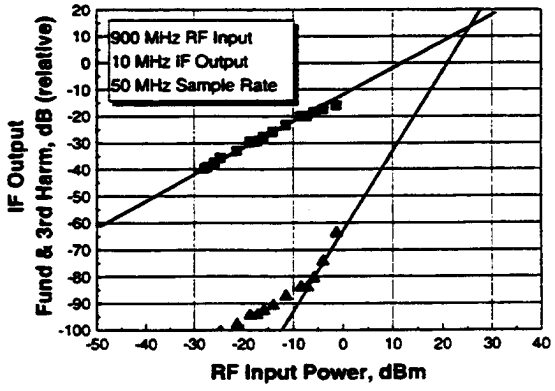


Fig.6. Measured mixer nonlinearity as third-order intercept (IP3). Discrete-time analog signal was directly applied to spectrum analyzer. Small inputs limited by noise, higher inputs by distortion in signal source.

Isolated and Coupled Superquadric Loop Antennas for Mobile Communications Applications

M.A. Jensen and Y. Rahmat-Samii

Proceedings of the Third International Mobile Satellite Conference
pp. 581-586
June 1993

Isolated and Coupled Superquadric Loop Antennas for Mobile Communications Applications

Michael A. Jensen and Yahya Rahmat-Samii
Department of Electrical Engineering
University of California, Los Angeles
405 Hilgard Avenue
Los Angeles, CA 90024-1594, USA
Telephone: (310) 206-3847
Fax: (310) 206-3847

ABSTRACT

This work provides an investigation of the performance of loop antennas for use in mobile communications applications. The analysis tools developed allow for high flexibility by representing the loop antenna as a superquadric curve, which includes the case of circular, elliptical, and rectangular loops. The antenna may be in an isolated environment, located above an infinite ground plane, or placed near a finite conducting plate or box. In cases where coupled loops are used, the two loops may have arbitrary relative positions and orientations. Several design examples are included to illustrate the versatility of the analysis capabilities. The performance of coupled loops arranged in a diversity scheme is also evaluated, and it is found that high diversity gain can be achieved even when the antennas are closely spaced.

INTRODUCTION

Circular and non-circular loop antennas, with shapes governed by packaging considerations, often prove to be appropriate, low-profile radiators for mobile communications devices. In some instances, new efforts to combat the effects of multipath fading without requiring increased bandwidth has motivated the use of multiple elements arranged in a diversity configuration on a single transceiver [1]. In the design of such antennas, it is important to understand the effects of loop geometry and mutual coupling on the antenna impedance, radiation characteristics, and diversity performance. This paper presents the results of two sophisticated analysis tools developed for this purpose:

(a) a Galerkin moment method algorithm for loops which are isolated or located near an infinite ground plane;
(b) a finite difference time domain (FDTD) technique for loops which are isolated or placed near finite-sized conducting objects.

ANTENNA GEOMETRY

To allow characterization of a wide variety of antenna geometries with one unified formulation, the loops are modeled as superquadric curves. This geometry is a closed loop which satisfies the equation

$$|x/a|^{\nu} + |y/b|^{\nu} = 1 \qquad (1)$$

where a and b are the semi-axes in the x and y directions respectively and ν is a "squareness parameter" which controls the variation of the loop radius of curvature. The configuration is illustrated in Fig. 1 for $\nu = 2$, 3, and 10 and an aspect ratio of $b/a = 2$. As can be seen, variation of the values of a, b, and ν allows considerable flexibility in modeling many practical antenna

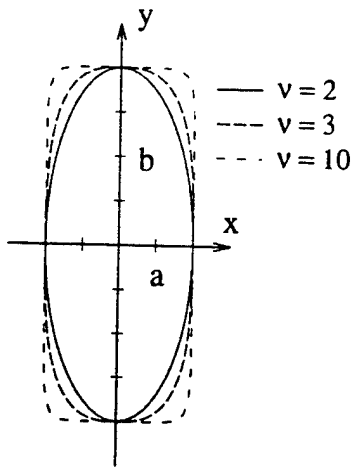


Figure 1: Superquadric geometry for $\nu = 2$, 3, and 10 with an aspect ratio of $b/a = 2$.

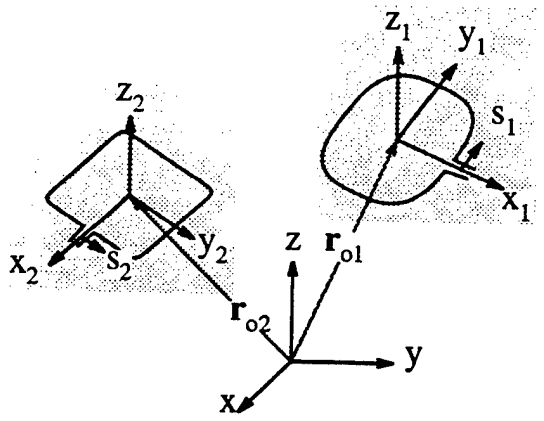


Figure 2: Geometry of two coupled loop antennas showing coordinates.

configurations. This flexibility is very important from the viewpoint of antenna packaging considerations.

For coupled loop geometries, the two superquadric antennas may have arbitrary positions and orientations and possibly different geometries, as represented in Fig. 2. Each loop is situated in its own coordinate system which may have an arbitrary position and orientation (described using Eulerian angles) with respect to the reference coordinate system.

FORMULATION

Moment Method Analysis

The moment method analysis of the coupled loop configuration makes use of a parametric expression for the superquadric curve in a coupled form of an electric field integral equation (EFIE) for thin wires. Use of this parametrization allows integration to occur on the curved loop contour rather than on the commonly-used piecewise linear representation of the curve, resulting in a more computationally efficient algorithm. Piecewise sinusoidal subsectional basis and weighting functions are used in a Galerkin form of the moment method to compute the axial current distribution along the loop. This current is then used to compute the antenna radiation pattern, directivity, and input impedance. The formulation is extended to analyze loops placed near an infinite ground plane through modification of the Green's function in the EFIE to account for the loop image. Both delta gap and magnetic frill type source models are used as excitation schemes to allow investigation of different feeding scenarios

FDTD Analysis

In order to fully evaluate the performance of loop antennas on a small box such as might be used for a hand-held radio transceiver, the effects of this conducting case must be included in the analysis. In order to investigate this configuration, the finite-difference time-domain (FDTD) algorithm is used with Yee's cubical cells and a second order absorbing boundary condition at the outer grid truncation surface. A special subcell method is used to properly account for the finite size of wires on the antenna radiation and impedance characteristics. By using properly shaped excitation functions for the antenna feed, the antenna behavior over a wide frequency

band may be determined with this time-domain formulation. An example of a design based upon these computations is provided at the end of this paper.

DIVERSITY

One of the objectives of this work is to determine the performance of the coupled superquadric loop antennas when used in a diversity scheme for a mobile communications system. In this application, we are interested in the use of space, angle, and polarization diversity to combat the effects of short-term or Rayleigh-type fading in a multipath environment. Antenna diversity scenarios in which multiple elements are used at the receiver to reduce the effects of fading are becoming more predominant as communications systems demand increased signal quality and reliability without consuming additional use of the available frequency spectrum.

A quantitative figure of merit for the performance of an antenna diversity configuration is the envelope correlation coefficient for the signals received by two different elements. In essence, this quantity provides a measure of the "similarity" of the two signals. For cases where the incident multipath field is assumed to arrive from the horizontal plane only it can be shown [1] that the envelope correlation coefficient for two antennas may be computed from the equation

$$\rho_e = \frac{\left|\int_0^{2\pi} \vec{E}_1(\pi/2,\phi)\cdot\vec{E}_2^*(\pi/2,\phi)d\phi\right|^2}{\int_0^{2\pi}|\vec{E}_1(\pi/2,\phi)|^2 d\phi \int_0^{2\pi}|\vec{E}_2(\pi/2,\phi)|^2 d\phi} \quad (2)$$

where $\vec{E}_1$ and $\vec{E}_2$ are the vector patterns associated with each of the coupled loop antennas. Generally, a value of ρ_e less than about 0.7 provides acceptable diversity returns.

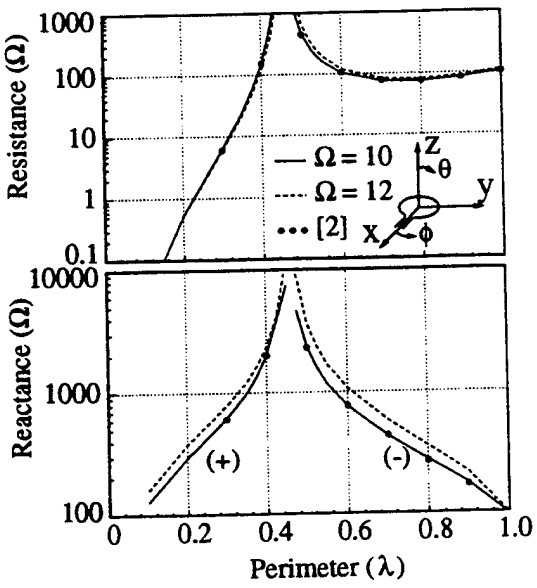


Figure 3: Input impedance versus perimeter for a single circular loop for two values of Ω. The dots are values taken from [2].

EXAMPLES

In the following examples, the parameter $\Omega = 2\ln(P/r_w)$ is used as a measure of the wire size where P is the loop perimeter and r_w is the wire radius. Fig. 3 shows the input impedance versus loop circumference for a single circular loop ($\nu = 2$, $b/a = 1$) for two values of Ω. A magnetic frill source model configured to have the dimensions of a 50Ω coaxial feeding line is used for the excitation. This plot shows that the loop antenna exhibits reasonable input impedance values for circumferences larger than about 0.7λ. The poor impedance behavior occurring for small loop circumferences can be substantially improved using proper loading techniques, thereby extending the loop to applications where the antenna size may be limited by spatial considerations. The dots in the figure correspond to data computed using a Fourier series representation for the current distribution [2]. Clearly, excellent agreement exists between the two sets of data.

An example of the use of the moment

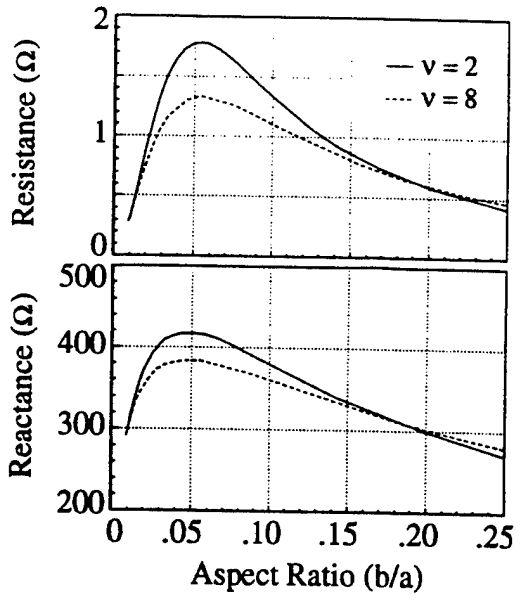


Figure 4: Input impedance versus aspect ratio b/a for a 0.25λ loop with $\Omega = 10$ for two values of ν.

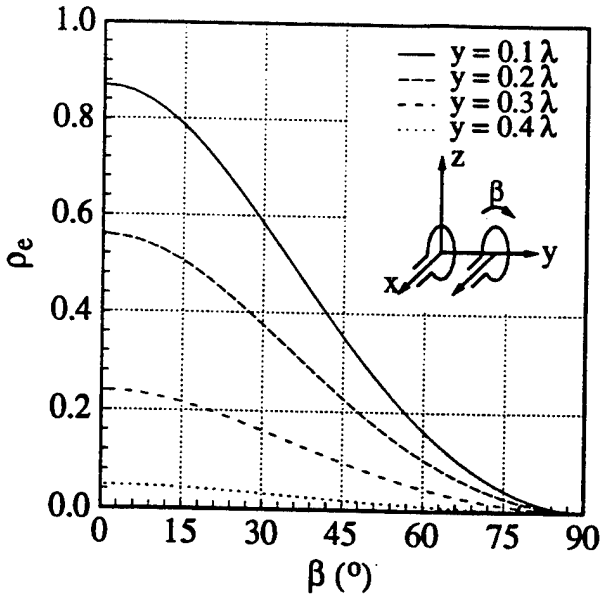


Figure 5: Envelope correlation coefficient versus rotation for 0.25λ loops ($\nu = 10$, $b/a = 1$) for various separation distances.

method to determine the effects of geometry on the loop antenna performance is illustrated in Fig. 4. This plot shows the input impedance of a 0.25λ loop with $\Omega = 10$ as the aspect ratio b/a is varied. Results are shown for $\nu = 2$ and 8. From this figure, it can be seen that the input impedance varies noticeably with loop squareness. This data can be very useful in the design of loop antennas

A key feature of loop antennas is that they may be configured to achieve high diversity performance when used in a multipath fading environment. For example, Fig. 5 illustrates the variation of the envelope correlation coefficient as a function of antenna orientation for two 0.25λ loops with $b/a = 1$ and $\nu = 10$ for several values of loop separation y. In this example, one loop is held stationary while the second is rotated about its x axis as shown in the figure inset. As can be seen from Fig. 5, low correlation coefficient values can be obtained even for small antenna spacings.

The high diversity performance of crossed loops leads to the possibility of a design such as that depicted in the inset of Fig. 6. The centroid of this antenna configuration is placed 0.2λ above an infinite ground plane to represent the scenario where the antenna is mounted on a car or other vehicle. Each loop is 0.8λ in perimeter with $\nu = 5$ and $b/a = 1$. The radiation patterns for these coupled loops normalized to the antenna directivity are shown in Fig. 6 for the principal and horizontal plane cuts. The antennas are fed 90° out of phase which results in the symmetry in the pattern in the xy plane. If proper signal combining is used with this antenna geometry, high diversity gain can be achieved.

An example of the flexibility of the FDTD methodology to predict the performance of a strip loop placed upon a hand-held transceiver case appears in Fig. 7. The geometry of the handset/antenna system is illustrated in Fig. 7(a). Fig. 7(b) demonstrates the wideband impedance behavior for this configuration At lower frequencies, the impedance varies rapidly with frequency which results in challenging matching requirements if broadband

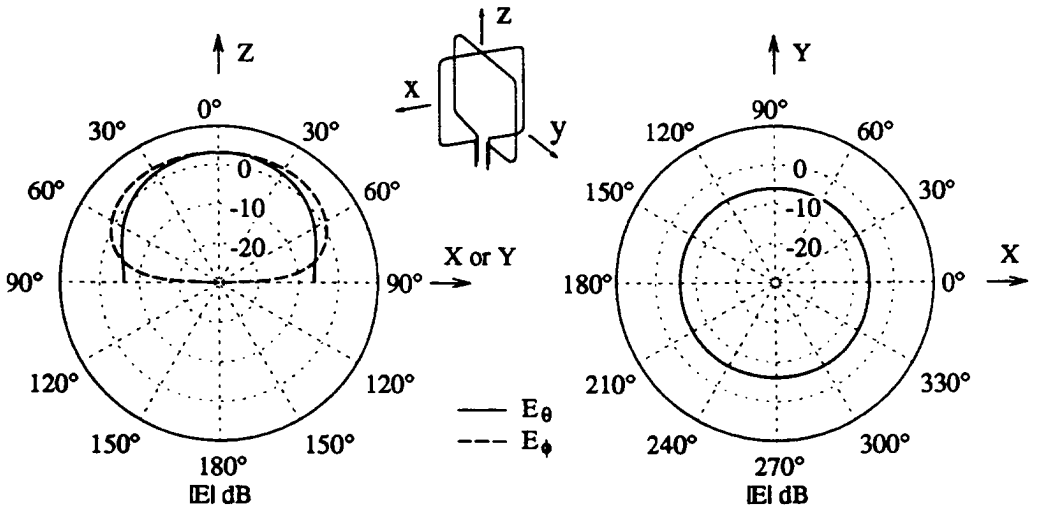


Figure 6: Directivity patterns in dB for two crossed 0.8λ superquadric loops ($\nu = 5$, $b/a = 1$) located $z_o = 0.2\lambda$ from an infinite ground plane.

performance is necessary. However, for narrowband applications, reasonable impedance values occur between the resonance peaks. For higher frequencies, the slow impedance variation with frequency allows for much wider bands of operation. The directivity patterns for the loop are provided in Fig. 7(c) for a frequency of $f = 915$MHz at which point the input impedance has a value of $53 - j246\Omega$. The asymmetries in the xz plane pattern arises from the fact that the loop is not centered on the box in the x direction to allow room for the feeding circuitry.

An example of using loading to improve the impedance characteristics of small loop antennas is provided in Fig. 8 which shows the input impedance versus frequency for a rectangular loop antenna loaded with an open circuit opposite the feed point. The loop has dimensions $a = 0.86$ cm and $b = 2.56$ cm, with a wire radius of 0.75 mm. This broadband data obtained with the FDTD methodology shows very well behaved impedance characteristics for low frequency operation, especially near 1GHz where the loop is near $\lambda/2$ in perimeter. Such an antenna may be appropriate for hand-held transceiver applications.

CONCLUSIONS

In this paper we have demonstrated the utility of newly developed computational tools based on both moment method and FDTD algorithms in the analysis of loop antennas for mobile communications applications. The analysis has been focussed on loops described by superquadric curves to allow characterization of a large number of geometries for isolated and coupled loop configurations. The diversity performance of coupled loops has been discussed and it was shown that high diversity gain is possible even for closely spaced antennas. Several key design examples were presented to illustrate the flexibility of the simulation capabilities. Naturally, these analysis tools can be applied to the characterization of numerous other possible configurations for mobile communications antennas.

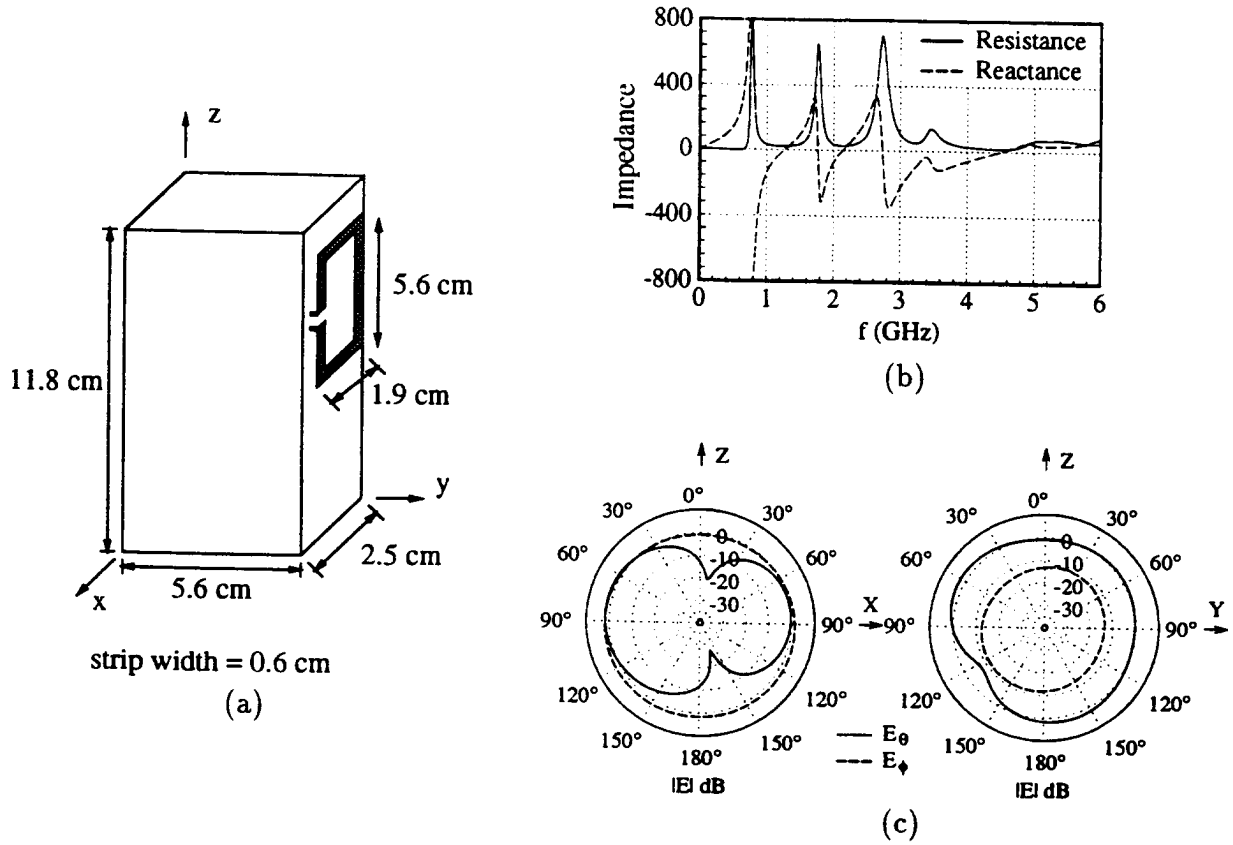


Figure 7: FDTD analysis of a strip loop on a hand-held transceiver: (a) transceiver geometry; (b) input impedance versus frequency; (c) directivity patterns at $f = 915$MHz.

Acknowledgements. This work is funded by DARPA under contract #DAAB07-92-R-C977. M. Jensen's work is also supported under a National Science Foundation Graduate Fellowship.

REFERENCES

[1] R. Vaughan and J. Andersen, "Antenna diversity in mobile communications," *IEEE Trans. Veh. Technol.*, **VT-36**, pp. 149–172, 1987.

[2] G. Zhou and G. Smith, "An accurate theoretical model for the thin-wire circular half-loop antenna," *IEEE Trans. Antennas Propag.*, **39**, pp. 1167–1177, 1991.

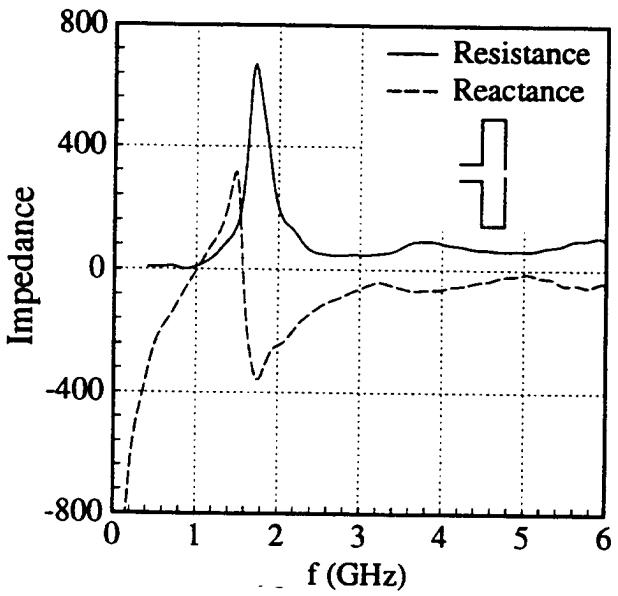


Figure 8: FDTD result of the input impedance versus frequency for a rectangular loop with $a = 0.86$ cm, $b = 2.56cm$, and $r_w = 0.75$ mm loaded with an open circuit.

Characteristics of Superquadric Loop antennas for Mobile Communications Applications

M.A. Jensen and Y. Rahmat-Samii

Proceedings of the IEEE Antennas and Propagation Society International Symposium
June 1993

CHARACTERISTICS OF SUPERQUADRIC LOOP ANTENNAS FOR MOBILE COMMUNICATIONS APPLICATIONS

Michael A. Jensen* and Yahya Rahmat-Samii
Department of Electrical Engineering
University of California, Los Angeles
Los Angeles, CA 90024-1594

I. INTRODUCTION In mobile communications systems, circular or non-circular loop antennas are often used either as single elements or in a multi-element array or diversity scheme. In the design of such antennas, it is important to understand the effects of loop geometry and coupling effects on the input impedance, radiation pattern, and diversity performance of the antenna configuration.

II. FORMULATION This paper presents the results of a unified formulation capable of modeling many different loop geometries, including circular, square, and folded dipole configurations, through use of the superquadric curve. The methodology allows analysis of a single superquadric element or of two loops with arbitrary relative position and orientation and possibly different geometries. A Galerkin type moment method with piecewise sinusoidal basis and weighting functions is used to solve a coupled form of the electric field integral equation (EFIE) for thin wires. The current obtained from this solution is subsequently used to compute the directivity, input impedance, and correlation coefficient (diversity performance) for the antenna configuration. Singularities in the integrand for coincident source and observation points are treated using a thin wire approximation. Both delta gap and magnetic frill source models are implemented in the formulation. To enhance the accuracy and efficiency of the methodology, curved rather than the commonly employed linear wire segments are used in the course of the implementation of the moment method.

III. SUPERQUADRIC LOOP A description of the superquadric curve is given by the equation

$$|x/a|^{\nu} + |y/b|^{\nu} = 1 \tag{1}$$

where a and b are the semi-axes in the x and y directions respectively and ν is a "squareness parameter" which controls the behavior of the loop radius of curvature. Fig. 1 illustrates the superquadric geometry for $\nu = 2$, 3, and 10 and an aspect ratio of $b/a = 2$. It is evident that the superquadric loop allows modeling of numerous antenna configurations through variation of the shape parameters a, b, and ν. A parametric representation of this equation is implemented to allow use of curved wire segments in the moment method analysis. Fig. 2 presents the geometry for the two coupled superquadric antennas denoted as loops 1 and 2. The two loops may have arbitrary positions and orientations and different geometries.

0-7803-1246-5/93/$3.00 © 1993 IEEE.

IV. EXAMPLES Unless otherwise indicated, a wire radius r_w is used such that

$$\Omega = 2 \ln \frac{P}{r_w} = 10 \tag{2}$$

is satisfied, where P is the loop perimeter. Fig. 3 shows the input impedance versus loop circumference for a single circular loop ($\nu = 2$, $b/a = 1$) for two values of Ω. A magnetic frill source model configured to have the dimensions of a 50Ω coaxial feeding line is used for the excitation. An interesting feature in this plot is the high impedance values occurring for loop circumferences near $\lambda/2$. The dots in the figure correspond to data computed using a Fourier series representation for the current distribution [1]. Clearly, excellent agreement exists between the two sets of data. The directivity pattern for the $\hat{\theta}$ component of a folded dipole with $r_w = 0.005\lambda$, $a = 6.5r_w$, $b = 0.25\lambda$, and $\nu = 50$ appears in Fig. 4. This $\lambda/2$ antenna is fed using a 0.0065λ delta gap source. This is the familiar 'doughnut' pattern associated with dipole antennas.

Fig. 5(b) provides the variation in input impedance versus aspect ratio b/a for a nearly rectangular ($\nu = 50$) delta gap fed 0.2λ loop located $c = 0.01\lambda$ from an infinite ground plane for three source positions as illustrated in Fig. 5(a). The computation is performed using coupled loops oriented such that they represent a loop plus its image. Although source position does have an impact on the loop impedance, it is clear that the aspect ratio plays a more dominant role in determining the impedance value. Fig. 6 shows the variation in input impedance versus loop squareness ν for two concentric loops each with $b/a = 1$. The larger and smaller loops are each 0.25λ at their respective operating frequencies of f_o and $3f_o$. To obtain this plot, each loop is in turn excited with a delta gap source while the other is short circuited. The solid line represents the impedance variation for a single isolated 0.25λ loop.

An important quantity in the design of mobile-based loops operating in a multipath-fading environment is the envelope correlation coefficient ρ_e for signals received by two closely-spaced antennas. To obtain this measure of diversity performance, the current distribution for each of the coupled loops is used to compute the vector radiation patterns $\vec{E}_1$ and $\vec{E}_2$ associated with each antenna. The coefficient ρ_e is then computed using a simple integration of the inner product of the patterns [2]. Fig. 7 illustrates the variation of the envelope correlation coefficient as a function of antenna orientation for two 0.25λ loops with $b/a = 1$ and $\nu = 8$ for several values of loop separation y. In this example, one loop is held stationary while the second is rotated about its x axis as shown in the figure inset. As can be seen from Fig. 7, low correlation coefficient values can be obtained even for small antenna spacings. This result is important in the design of diversity antennas for use in miniature portable communications equipment.

Acknowledgements. This work is funded by DARPA Contract #DAAB07-92-R-C977. M. Jensen's work is also supported under a National Science Foundation Graduate Fellowship.

REFERENCES

[1] G. Zhou and G. Smith, "An accurate theoretical model for the thin-wire circular half-loop antenna," *IEEE Trans. Antennas Propag.*, **39**, pp. 1167–1177, 1991.

[2] R. Vaughan and J. Andersen, "Antenna diversity in mobile communications," *IEEE Trans. Veh. Technol.*, **VT-36**, pp. 149–172, 1987.

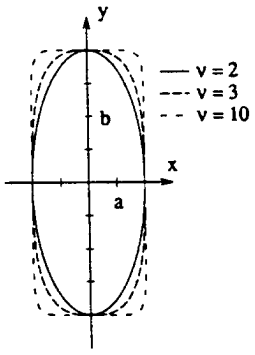


Figure 1: Superquadric geometry for an aspect ratio of $b/a = 2$.

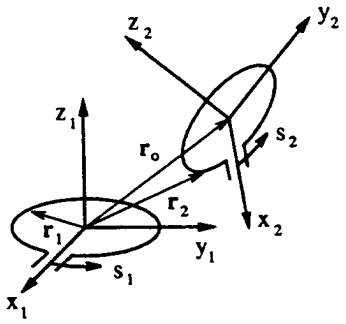


Figure 2: Geometry of two coupled loop antennas showing coordinates.

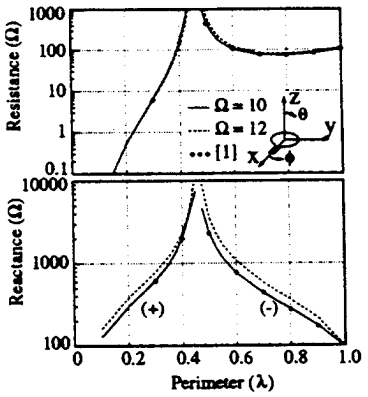


Figure 3: Input impedance versus perimeter for a single circular loop for two values of Ω.

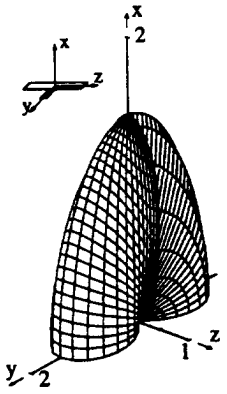


Figure 4: Directivity pattern in one hemisphere for a $\lambda/2$ folded dipole with $\nu = 50$, $r_w = 0.0005\lambda$, $a = 6.5r_w$, and $b = 0.25\lambda$.

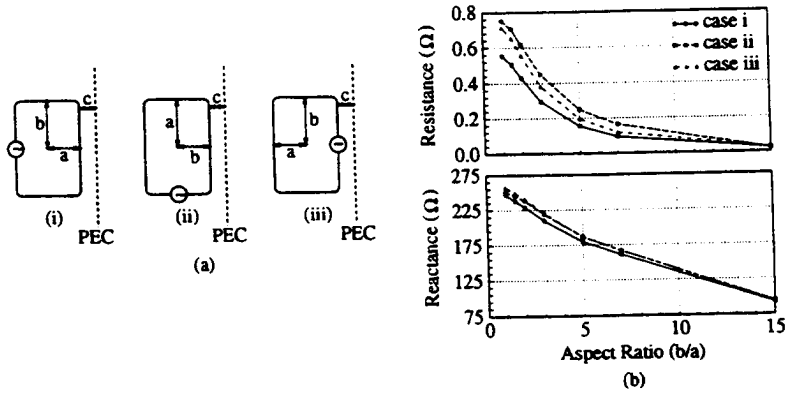


Figure 5: Superquadric loops ($\nu = 50$, $P = 0.2\lambda$, $c = 0.01\lambda$) near an infinite ground plane: (a) geometry and (b) input impedance versus aspect ratio b/a.

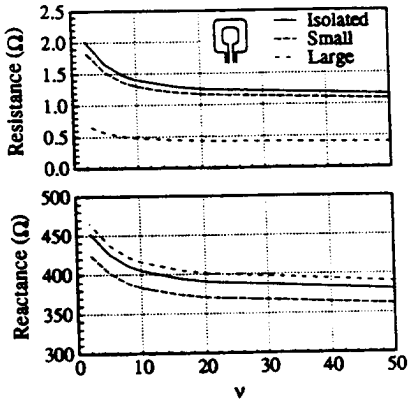


Figure 6: Input impedance versus squareness parameter ν for concentric loops with $b/a = 1$ operating at f_o and $3f_o$ compared with that of an isolated loop.

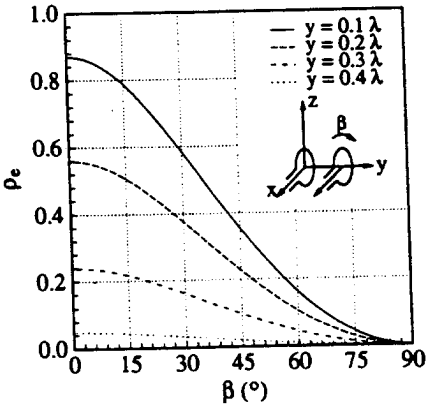


Figure 7: Envelope correlation coefficient versus antenna orientation for two 0.25λ loops with $b/a = 1$ and $\nu = 8$. One loop is held stationary while the second is rotated.

FDTD Analysis of PIFA Diversity Antennas on a Handheld Transceiver Unit

M.A. Jensen and Y. Rahmat-Samii

Proceedings of the IEEE Antennas and Propagation Society International Symposium
June 1993

FDTD ANALYSIS OF PIFA DIVERSITY ANTENNAS ON A HAND-HELD TRANSCEIVER UNIT

Michael A. Jensen* and Yahya Rahmat-Samii
Department of Electrical Engineering
University of California, Los Angeles
Los Angeles, CA 90024-1594

I. INTRODUCTION Because of its high efficiency and bandwidth, the monopole antenna is generally used for portable communications transceivers. However, increased efforts to miniaturize hand-held communication devices has led to the development of low-profile elements such as the planar inverted-F antenna (PIFA) [1]. The small, unobtrusive nature of these antennas allows multiple elements to be placed on the handset in a diversity scheme to combat short-term fading. Furthermore, they exhibit an improved cross-polarization response which also proves beneficial in a multipath fading environment.

To properly design a diversity antenna consisting of two coupled PIFA elements on a conducting transceiver body, it is first necessary to characterize the antenna behavior using an analysis tool such as the finite-difference time-domain (FDTD) method [2]. In this paper, the FDTD technique is used to investigate the radiation patterns, polarization characteristics, and correlation coefficient for a diversity antenna employing two PIFA elements located on a rectangular conducting box. A second order absorbing boundary condition is used to truncate the computational volume at a distance of 12 cells away from the antenna/handset configuration. A wire subcell method is employed to properly incorporate the effects of the finite size of the feeding wire.

II. GEOMETRY The PIFA element with its accompanying dimensions are illustrated in Fig. 1. The antenna is fed using a coaxial cable emerging through the ground plane below the element. Numerically, this feeding situation is simulated by introducing sources to provide radial electric field components at a source plane within the coax as shown in the figure. This scenario is similar to the magnetic frill excitation model often used in moment method analysis. Reflections of backscattered fields by the source model are avoided by implementing the excitation as a current source and terminating the coax with a first-order absorbing boundary condition several cells behind the source plane.

Fig. 2 depicts the rectangular conducting box with candidate element positions denoted as 1-4. All four elements are centered in the x-direction on the box. For the case of elements 1 and 2, the feeding coaxial line is located a distance t_z from the top of the box at $z = c$. Similarly, the feeding coax for both elements 3 and 4 is located a distance t_y from the box edge at $y = b$. The dimensions used for the PIFA and the box are detailed in Table 1. The inner conductor of the coax has a radius of 0.0025λ.

0-7803-1246-5/93/$3.00 © 1993 IEEE.

III. RESULTS To determine the radiation pattern and polarization characteristics of the PIFA elements on the conducting box, the FDTD method is applied with cell sizes of $\lambda/100$. The sinusoidal input is allowed to cycle through 7 periods to settle into steady state, after which the tangential field components on a surface enclosing the handset are collected during an additional cycle. Because the FDTD algorithm produces field components which are spatially offset from each other, a simple spatial average is taken such that all field components lie on a single surface. An FFT is used to obtain the amplitude and phase information of the fields at each sample point and a frequency domain near-to-far field transformation is implemented to compute the radiation patterns.

Fig. 3 illustrates the radiation patterns for several of the elements of Fig. 2. Each pair of patterns ($\phi = 0°$ and $\phi = 90°$ planes) is normalized such that the maximum value is 0 dB. Fig. 3(a) and Fig. 3(b) show the amplitude patterns for elements 1 and 3 respectively. Fig. 3(c) illustrates the pattern when elements 1 and 2 are operating simultaneously, and Fig. 3(d) shows a similar case for elements 3 and 4. As can be seen from these results, the use of two elements produces deeps nulls in the overall radiation pattern of the antenna. This is especially noted in the $\phi = 0°$ pattern of Fig. 3(c), where the $\hat{\phi}$ component of the field essentially vanishes. The plots further illustrate the difference in patterns associated with different element locations. This pattern diversity can be used to advantage to combat short-term fading in multipath environments.

The envelope correlation coefficient ρ_e provides a measure of the effectiveness of antennas arranged in a diversity scheme. For a Rayleigh distributed incoming signal, this coefficient has a range $0 \leq \rho_e \leq 1$. It is generally accepted that a correlation coefficient $\rho_e \leq 0.7$ is sufficiently low to provide acceptable diversity gain. If it is assumed that the incoming multipath waves exist only in the horizontal plane and that either polarization is equally likely to occur, it can be shown that the envelope correlation coefficient can be computed from the relation [3]

$$\rho_e = \frac{\left| \int_0^{2\pi} \vec{E}_1(\pi/2, \phi) \cdot \vec{E}_2^*(\pi/2, \phi) d\phi \right|^2}{\int_0^{2\pi} |\vec{E}_1(\pi/2, \phi)|^2 d\phi \int_0^{2\pi} |\vec{E}_2(\pi/2, \phi)|^2 d\phi} \tag{1}$$

where $\vec{E}_1$ and $\vec{E}_2$ are the patterns of the two antennas in the system.

The diversity performance for several combinations of two elements is provided in Table 2. The values for ρ_e are obtained by first computing the patterns for the individual elements on the transceiver body and using these patterns in (1). The high value of ρ_e for the element pair 3-4 can be explained by realizing that vertical separation does not provide spatial diversity for waves incident from the horizontal plane. For the other cases, the combination of spatial, polarization, and angle diversity acts to provide very low correlation coefficients.

Acknowledgements. This work is funded by DARPA Contract #DAAB07-92-R-C977. M. Jensen's work is also supported under a National Science Foundation Graduate Fellowship.

REFERENCES

[1] T. Taga and K. Tsunekawa, "A built-in diversity antenna for 800MHz band portable radio units," *IEEE Trans. Selected Areas Commun.*, **SAC-5**, pp. 921–929, 1987.

[2] A. Taflove and M. E. Brodwin, "Numerical solution of steady- state electromagnetic scattering problems using the time-dependent Maxwell's equations," *IEEE Trans. Microwave Theory Tech.*, **MTT-23**, pp. 623–630, 1975.

[3] R. Vaughan and J. Andersen, "Antenna diversity in mobile communications," *IEEE Trans. Veh. Technol.*, **VT-36**, pp. 149–172, 1987.

Table 1: Dimensions used in the analysis of the PIFA on a conducting box.

Dimension	*Size* (λ)
a	0.12
b	0.24
c	0.38
h	0.04
w	0.10
d	0.14
t_y	0.05
t_z	0.02
s	0.02

Table 2: Envelope correlation coefficient for combinations of elements shown in Fig. 2.

Element Pair	ρ_e
1-2	0.01606
1-3	0.11374
1-4	0.16182
2-3	0.01653
2-4	0.03923
3-4	0.89067

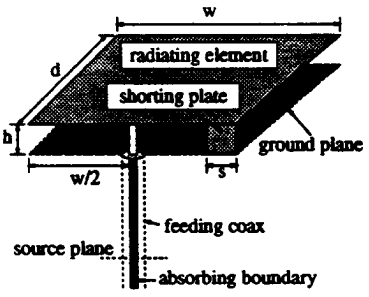


Figure 1: PIFA geometry and feeding arrangement.

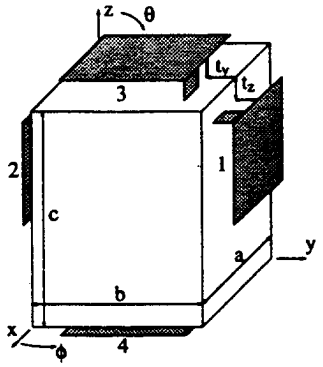


Figure 2: Geometry of the conducting box with potential locations of PIFA elements.

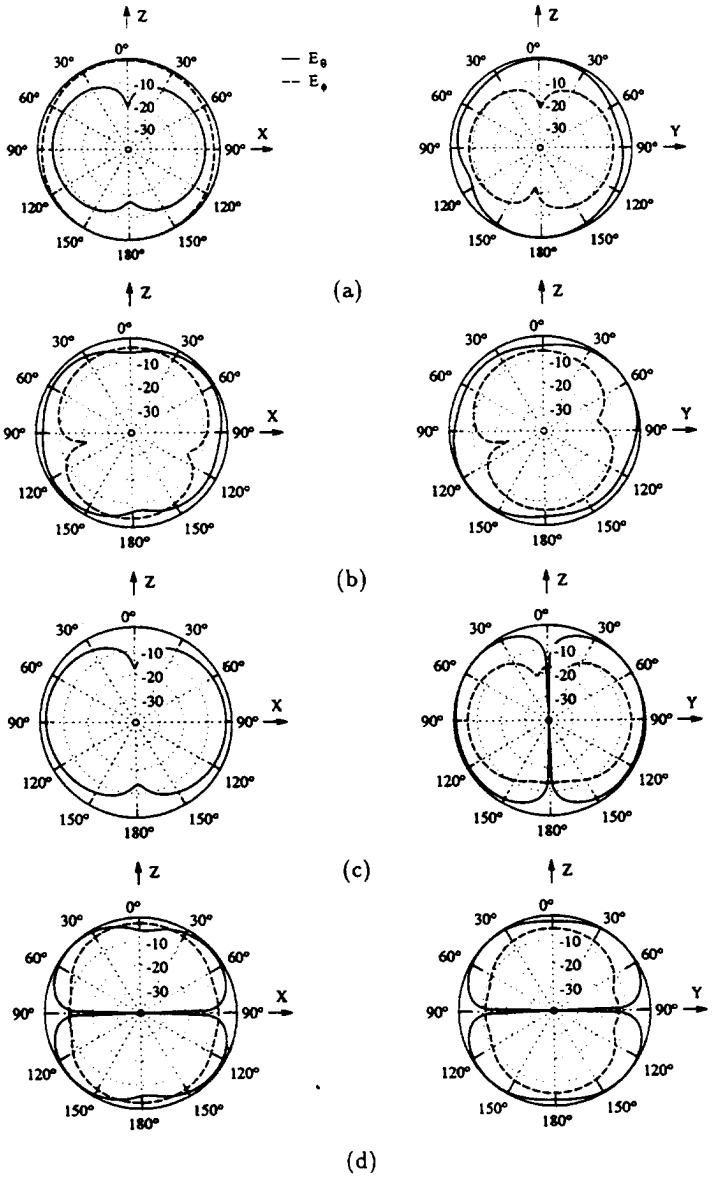


Figure 3: Radiation patterns in dB for several combinations of elements in Fig. 2: (a) element 1; (b) element 3; (c) elements 1 and 2; (d) elements 3 and 4.

Large Suspended Inductors on Silicon and Their Use in a 2-μm CMOS RF Amplifier

J.Y.-C. Chang, A.A. Abidi, and M. Gaitan

IEEE Electron Device Letters
vol. 14, no. 5, pp. 246-248
May 1993

Large Suspended Inductors on Silicon and Their Use in a 2-μm CMOS RF Amplifier

J. Y.-C. Chang, *Student Member, IEEE,* Asad A. Abidi, *Member, IEEE,* and Michael Gaitan, *Member, IEEE*

Abstract—**Large spiral inductors encased in oxide over silicon are shown to operate beyond the UHF band when the capacitance and loss resistance are dramatically reduced by selective removal of the underlying substrate. Using a 100-nH inductor whose self-resonance lies at 3 GHz, a balanced tuned amplifier with a gain of 14 dB centered at 770 MHz has been implemented in a standard digital 2-μm CMOS IC process. The core amplifier noise figure is 6 dB, and the power dissipation is 7 mW from a 3-V supply.**

I. Introduction

THE growing needs for miniature wireless communication in the 1-GHz band have prompted interest in monolithic RF amplifiers in silicon. Modern bipolar transistors and FET's certainly have a high enough f_T to provide gain in the required narrow frequency band at these frequencies; the challenge, interestingly enough, is in the difficult fabrication of monolithic passive components. An RF amplifier employs a tuned load, to act as a secondary filter for out of band signals and noise following preselection at the antenna, but most importantly as a means to obtain gain which may be as large as that available at dc by using LC resonance to null out device and parasitic capacitances at the center frequency.

Efforts to fabricate large value spiral inductors on silicon substrates in the 1960's [1] led to the conclusion that the self-resonance caused by parasitic capacitance of these structures to the substrate would limit their use at high frequencies, and the series spreading resistance in the lossy substrate their quality factor (Q). Spiral inductors were revisited when the fabrication of tuned amplifiers became common on the semi-insulating GaAs substrate, where these limitations no longer exist. Most GaAs circuits required inductances of only a few nanohenries.

The work reported here was motivated by the need to develop a silicon *low-power* RF amplifier in the 800- to 900-MHz frequency band. A low-power amplifier must necessarily operate at a high impedance level, and therefore requires large inductors in the tuned load. Previous work [2] in a 0.3-μm GaAs MESFET technology has demonstrated a monolithic RF amplifier tuned to 1.6 GHz with a conventional on-chip inductor of about 12 nH. We set ourselves the greater challenge of obtaining a similar performance in the 700- to 900-MHz band with a 2-μm digital silicon CMOS technology. The lower frequency of operation means that the tuned load requires inductors as large as 100 nH. Lessons learned from past efforts in inductor fabrication on silicon suggested, however, that, short of a fundamental innovation, this was an impossible goal. How this challenge was met is described in the following sections.

Manuscript received February 12, 1993. This work was supported by DARPA, Rockwell International, and the State of California MICRO Program. In this report, commercial computer programs and instruments are identified to specify the procedure adequately. This does not imply recommendation or endorsement by NIST, nor that the program is the best available for the purpose.

J. Y.-C. Chang and A. A. Abidi are with the Integrated Circuits and Systems Laboratory, Electrical Engineering Department, University of California, Los Angeles, CA 90024-1594.

M. Gaitan is with the Device Technology Group, Semiconductor Electronics Division, National Institute of Standards and Technology, Gaithersburg, MD 20899.

IEEE Log Number 9208809.

II. A Suspended Inductor

Large-value inductors may be fabricated as aluminum spirals with many turns. As the inductance of the spiral is made larger, the capacitance to substrate increases, leading to a progressively *lower* frequency of self-resonance. Spiral inductors of 25 nH are found to self resonate at about 3 GHz on GaAs substrates [2] and on insulating sapphire substrates [3]. On the other hand, aluminum inductors only as large as 10 nH on standard silicon substrates will self resonate at 2 GHz [4], and furthermore the spreading resistance of the substrate will introduce a loss.

These characteristic problems of a silicon substrate may be overcome if the area under the inductor is made to appear locally insulating. This is most simply accomplished by selectively etching out the silicon, leaving the inductor encased in a suspended oxide layer attached at four corners to the rest of the silicon IC. There is a similarity between this technique and the practice in some GaAs technologies of suspending spiral inductors on air bridges [5], but as the typical gap under an air bridge is 3 μm, while removal of the substrate offers air gaps as large as 200 to 500 μm, inductors fabricated by our technique obtain a much lower capacitance to substrate.

A 100-nH inductor was designed using analytical formulas [6] as a 20-turn square spiral of 4-μm-wide lines of second-layer aluminum metal separated by 4-μm spaces, resulting in an outer dimension of 440 μm. Simulations on the SONNET® EM 3-D electromagnetic simulator showed that removal of the underlying substrate will cause the inductor self-resonance to move out from 800 MHz to 3 GHz (Fig. 1).

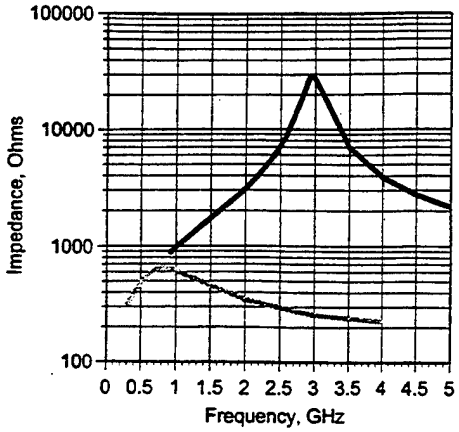


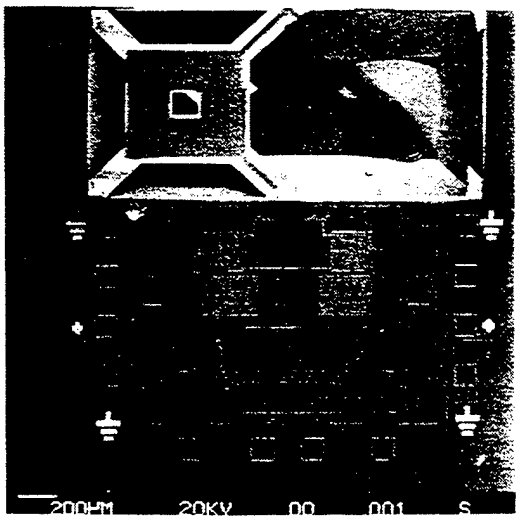
Fig. 1. Self-resonance frequency of a 100-nH inductor increases from 800 MHz to 3 GHz after removal of underlying substrate. This corresponds to a 14-fold reduction in parasitic capacitance. Impedance of inductor over silicon (gray) and over pit (black) obtained from 3-D electromagnetic simulations.

Accompanied by active circuits for an amplifier and buffers, this structure was fabricated through MOSIS as a standard n-well 2-μm CMOS IC. Using a previously described technique [7], the fabricated die were then subject to a selective EDP wet etch to remove the substrate under the inductors, while leaving the remaining circuits intact (Fig. 2). The starting areas for the etch were defined without use of an extra mask [7].

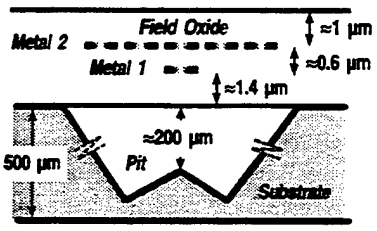
III. RF Amplifier and Experimental Results

Two identical inductors of nominal value 100 nH were used as the load of a balanced amplifier, tuned to a center frequency of 800 MHz by the junction capacitance of FET's built with 2-μm design rules (Fig. 3). The circuit attained high-frequency gain at low power at the expense of noise figure. Cascode stages were used to enhance the available gain. For ease of testing in a network analyzer, on-chip polysilicon 50-Ω resistors terminated the two inputs to ground. The outputs were buffered through common-source FET's to a 50-Ω environment off-chip. The buffer frequency response was separately measured to deembed the core amplifier.

Operating at a 3-V power supply, a peak gain of 14 dB centered at 770 MHz was obtained in the core amplifier with a 7-mW power dissipation. The measured frequency response conformed well to simulations (Fig. 4), although the center frequency was lower by 50 MHz owing to a larger FET junction capacitance than anticipated. Chips where the substrate was left intact under the inductor show a dramatically worse frequency response (Fig. 4). The measured noise figure of the core amplifier, excluding the noise contribution of the termination resistors, was 6 dB. As the substrate capacitance and loss in the suspended inductor are negligibly small, it was modeled for the purpose of circuit simulation by a simple *LCR* equivalent circuit, whose parameter values correspond very well to Cascade® probe measurements on a test inductor (Fig. 4). The large suspended structure is found to be mechanically very robust. No damage has yet been observed on an IC when it sustained impacts during the packaging procedure and while measurements were being taken.



(a)



(b)

Fig. 2. (a) SEM of RF amplifier after selective substrate etch. Inductor shown is suspended on oxide layer attached to substrate at four corners. Spiral fabricated as second-level aluminum, while contact from center brought out on first level. Second inductor has been manually removed to show pit. (b) Cross section of suspended inductor and substrate after etching.

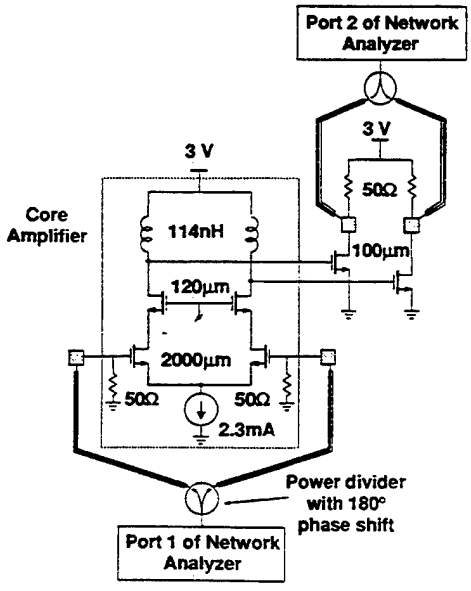


Fig. 3. Schematic of CMOS RF amplifier, including test arrangement. All FET's are 2-μm channel length. Bond pads define the chip periphery.

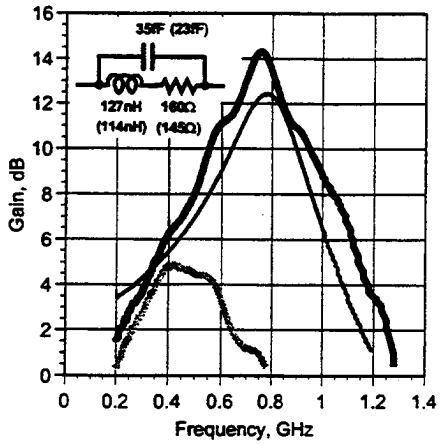


Fig. 4. Measured frequency response of core amplifier, after (black) and before (gray) substrate removal. Simulated response from SPICE (thin black) using a two-terminal *LCR* inductor model obtained from EM simulations. Measured parameter values in equivalent circuit (inset) compare favorably with values obtained from 3-D simulations (in parentheses).

IV. Conclusions

A use of selective etching is shown to obviate problems which were thought to plague all selective RF circuits on silicon substrates. As an added bonus, the etching requires no modifications to a standard digital CMOS IC process. The test vehicle is possibly the highest performance 2-μm CMOS amplifier reported to date, with a 14-dB gain at 770 MHz while requiring only a 3-V power supply from which it drains 7 mW.

References

[1] R. M. Warner, Ed., *Integrated Circuits: Design Principles and Fabrication*. New York: McGraw-Hill, 1965.

[2] Y. Imai, M. Tokumitsu, and A. Minakawa, "Design and performance of low current GaAs MMICs for *L*-band front-end applications," *IEEE Trans. Microwave Theory Tech.*, vol 39, pp. 209–215, Feb. 1991.

[3] M. Caulton, S. P. Knight, and D. A. Daly, "Hybrid integrated lumped element microwave amplifiers," *IEEE Trans. Electron Devices*, vol. ED-15, pp. 459–466, July 1968.

[4] N. M. Nguyen and R. G. Meyer, "Si IC-compatible inductors and *LC* passive filters," *IEEE J. Solid-State Circuits*, vol. 25, pp. 1028–1031, Aug. 1990.

[5] M. E. Goldfarb and V. K. Tripathi, "The effect of air bridge height on the propagation characteristics of microstrip," *IEEE Microwave Guided Wave Lett.*, vol. 1, pp. 273–274, Oct. 1991.

[6] H. M. Greenhouse, "Design of planar rectangular microelectronic inductors," *IEEE Trans. Parts, Hybrids, Packaging*, vol. PHP-10, pp. 101–109, June 1974.

[7] M. Parameswaran, H. P. Baltes, Lj. Rustic, A. C. Dhaded, and A. M. Robinson, "A new approach for the fabrication of micromachined structures," *Sensors and Actuators*, vol. 19, pp. 289–307, Sept. 1989.

A 750 MHz RF Amplifier in 2-mm CMOS

J.Y.-C. Chang and A.A. Abidi

1993 Symposium on VLSI Circuits Digest of Technical Papers
pp. 111-112
May 1993

13-1

A 750 MHz RF Amplifier in 2-μm CMOS

J. Y-C. Chang and A. A. Abidi

Integrated Circuits & Systems Laboratory
University of California
Los Angeles, CA 90024-1594

Low-power RF and IF components are essential building blocks for future portable wireless communicators. The conventional view holds that only GaAs or silicon bipolar technologies offer the required gain and low noise at an acceptably low power dissipation in an RF amplifier operating in the 900 MHz frequency band. A wireless transceiver will therefore consist of a mixture of technologies for RF, IF, and baseband signal processing. However, if the entire transceiver is integrated in a single technology, just the elimination of off-chip buffers between the RF and IF sections, say, can lead to considerable power savings. CMOS is the only viable technology for a digital transceiver. To this end, we describe here a 2-μm CMOS tuned amplifier affording 14 dB gain centered at 750 MHz, 6 dB noise figure, and 7 mW power dissipation from a 3 V supply. The design of this prototype will be straightforwardly re-centered at 900 MHz in a future version.

Fundamental Considerations

All RF amplifiers are tuned with an inductive load to obtain narrowband gain. Tuning transforms the baseband frequency response of an amplifier to a bandpass response. It also filters out of band noise introduced by the amplifier itself. The tuned load is, however, inadequate for channel selection in the receiver, so usually a highly selective, external passive filter precedes the receiver.

Maximum amplifier gain is set by the transistor $g_m r_{out}$ product, and by the loss resistance in the inductive load. The target noise figure, and the loading of the following stages, determine the power dissipation. As the amplifier noise is mainly generated in the channel conductance of the input FET, the noise figure specification will determine input g_m. This, in turn, sets the bias current, I, according to the FET I-V characteristic $g_m/I = 2/(V_{GS} - V_t)$, and thus the power dissipation. The lower the effective gate voltage bias of the input FET, the smaller the power dissipated to satisfy the noise specification. As $(V_{GS} - V_t) \rightarrow 2kT/q$, the transconductance of a FET approaches that of a bipolar transistor at the same operating current. An RF amplifier will operate perfectly well when biased at these low levels because it is subject to input signals of at most a few millivolts.

A high impedance load should be connected to the amplifying FET to realize a large voltage gain at RF. An advantage of large gain is that later mixer and IF stages will not degrade the receiver noise figure. In the tuned load, high impedance requires a large value inductor, low capacitance, and large parallel resistance. Use of an off-chip inductor is ruled out because the parasitic capacitances of the pads and package will be excessively large. On-chip inductors on silicon substrates, plagued by substrate capacitance and losses, are usually thought to be impractical at high frequencies. Although on-chip spiral inductors of 10 nH have previously been used in a 1 GHz bipolar amplifier [1], they are an order of magnitude smaller than what we seek.

Large Value Monolithic Inductor

We have developed a new technique for the fabrication of large value inductors usable at high frequencies on silicon substrates. The technique is particularly attractive because it relies exclusively on a simple post-process etch without any modification to a standard CMOS process [2]. Subsequent to the etch, a large spiral of interconnect metal encased in oxide, is suspended at four corners over a pit in the silicon substrate. Thus, the problems of substrate capacitance and loss both disappear (Fig.1). The only capacitance of any consequence now is between the coils of the spiral, and this is considerably smaller than capacitance to substrate.

A 110 nH spiral was designed on a 3D electromagnetic structural simulation program to first verify the feasibility of this concept. The spiral inductor was then fabricated in a foundry n-well CMOS process, and experimentally evaluated using on-chip s-parameter measurements from 0.3 to 3 GHz with a Cascade® probe. A remarkably good fit was found between simulation and measurements, which verify that the large suspended inductor is usable to 2.5 GHz before it self-resonates. A simple three component *LCR* equivalent circuit (Fig.2) of the spiral obtained from 3D electromagnetic simulations, and verified experimentally, is used for circuit simulations.

Amplifier Design

A familiar design style in baseband circuits is applied to this RF amplifier (Fig.3). The input and output of the amplifier are fully balanced, lending a large rejection of common-mode interference from other circuits sharing its substrate. The input FETs are 2 mm wide, and following the rationale outlined above, are biased at $(V_{GS} - V_t) = 150$ mV. A cascode FET pair of 120 μm width tapers the output to a small conductance. The tuned load consists of the 114 nH spiral inductor, the capacitances of the cascode and common-source output FETs, and the loss conductance across the inductor, which dominates the cascode output conductance. On-chip 50 Ω polysilicon resistors are added to the input stage for ease of testing. Owing to the negligible DC drop across the inductor, the NMOSFET stack in the amplifier comfortably biases at a power supply of 3 V.

The inductor is a 440 μm square structure, consisting of 20 turns of second-layer metal lines of 4 μm width with 4 μm pitch

Research supported by DARPA, Rockwell International and the State of California MICRO Program

(Fig.4). Negligible mutual coupling is expected between the two independent inductors laid out side by side in the differential amplifier.

Experimental Results

Bare amplifier die are mounted on Duroid® boards and wire bonded from the pads to copper striplines. A phase splitting power divider generates a balanced excitation to the amplifier, and by reciprocal action, another converts the balanced output to single-ended (Fig.5(a)). The gain of replica common-source buffers is separately measured to de-embed the true frequency response of the tuned amplifier. The actual gain is about 2 dB larger than simulated, and there is excellent agreement in the frequency response (Fig.5(b)). Measurements also show the dramatic degradation in the amplifier gain if a pit is *not* etched under the spiral. Owing to variations in junction and FET capacitance across 8 samples, there is a one-σ spread of 7.2 MHz on the mean amplifier center frequency of 750 MHz; however, this spread constitutes only 2.5% of the –3dB bandwidth of 270 MHz.

After removing the contribution of the on-chip termination resistors to the overall measured noise figure, the core amplifier noise figure is found to be 6 dB, which is about 0.5 dB larger than prediction assuming that the channel conductance of the input FETs is the sole contributor to noise. The termination resistor noise is discounted because in our application the RF amplifier will be physically mounted on the antenna, so there is no need to lower the system impedance to 50 Ω for matching or interconnection. The noise figure may be further lowered if required at the expense of greater power dissipation by scaling up the widths of all FETs and scaling down the inductor.

This work has reported the first CMOS RF amplifier, which to our knowledge is also the highest frequency amplifier in 2-μm CMOS amplifier to date, and the lowest power in any technology. A new fabrication technique is used to obviate many of the parasitic effects previously thought to plague monolithic inductors on silicon substrates.

1. Nguyen, N.M. and R.G. Meyer, "Si IC-compatible Inductors and LC Passive Filters", *IEEE J. of Solid State Circuits*, v. 25, pp. 1028-1031, August 1990.

2. Parameswaran, M., et al., "Micromachined Thermal Radiation Emitter from a Commercial CMOS Process", *IEEE Electron Device Letters*, v. 12, pp. 57-59, February 1991.

Fig.1. Impedance of 114 nH spiral inductor on silicon and over a pit, obtained from 3D electromagnetic simulations. Removal of underlying substrate substantially eliminates large capacitance and losses, and pushes self-resonance to 3 GHz.

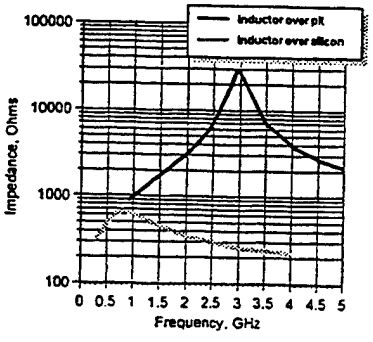


Fig.2. Equivalent circuit of suspended inductor. Measured parameter values are compared with values obtained from 3D simulations (in parantheses).

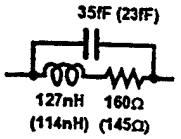
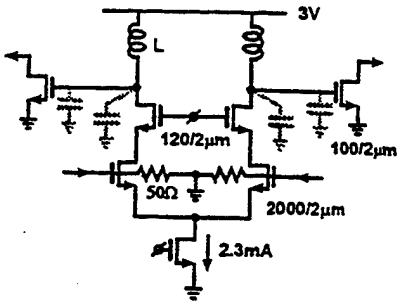


Fig.3. Circuit diagram of monolithic RF amplifier. Open drain ouput MOSFETs are biased off-chip.



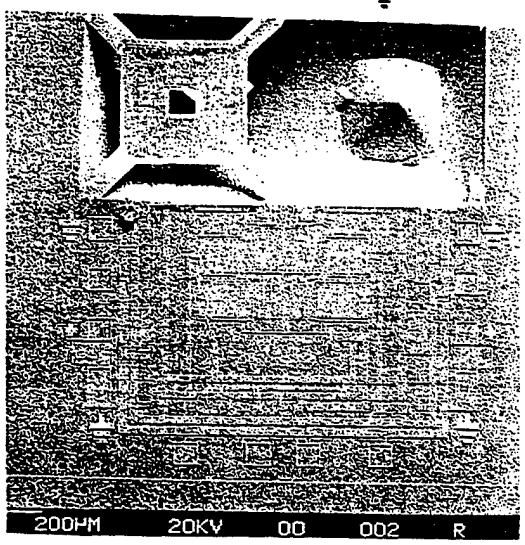
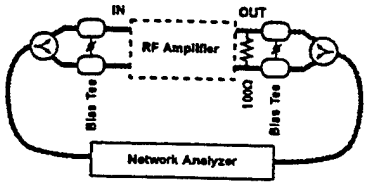


Fig.4. SEM of RF amplifier. One of the spiral inductors has been manually removed to show pit underneath.

(a)



(b)

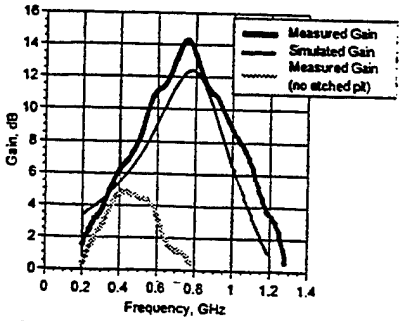


Fig.5. (a) Setup for network analyzer measurements. (b) Measured frequency response compares very well with simulations. Measurements also made on amplifier with no pit under inductor show advantage of suspended inductor.

Superquadric Loop Antennas

M.A. Jensen and Y. Rahmat-Samii

Eighth International Conference on Antennas and Propagation
pp. 103-106
March 1993

Eighth International Conference on

Antennas and Propagation

30 March - 2 April 1993

Organised by

the Electronics Division of the Institution of Electrical Engineers

in association with

EUREL
International Union of Radio Science (URSI)
Institute of Electronics, Information and Communication Engineers of Japan
Institution of Electrical and Electronic Engineers (UKRI Section)
Institution of Electrical and Electronic Engineers Inc (Antennas and Propagation Society)

Venue

Heriot- Watt University, UK

SUPERQUADRIC LOOP ANTENNAS

M A Jensen and Y Rahmat-Samii

University of California, Los Angeles, USA

INTRODUCTION

The increasing effort in miniaturization of mobile communications equipment has spurred the development of small, low-profile antennas suitable for implementation in portable devices. In many such applications, circular or non-circular loop antennas are used either as single elements or in a multi-element array or diversity scheme. It is important in the design of loop antennas to understand the effects of loop geometry and coupling effects on the input impedance, radiation pattern, and diversity performance of the antenna configuration.

In this paper a unified formulation is presented which allows modeling of a continuum of different loop geometries, including the well-known circular, square, and folded dipole configurations, through use of the superquadric curve. The analysis includes the case of two superquadric loops with arbitrary geometries, excitation, and relative orientation to allow determination of the effects of mutual coupling on the antenna characteristics. A Galerkin type moment method using piecewise sinusoidal basis and weighting functions is used to solve an electric field integral equation (EFIE) for thin wires. A novel parametrization for the superquadric curve is given which allows the use of curved wire segments, rather than the commonly employed linear segments, to construct the antenna geometry. The concept of antenna diversity is discussed and the diversity performance of coupled superquadric loop antennas is presented.

FORMULATION

Problem Geometry

The superquadric curve is the locus of points satisfying the equation

$$\left|\frac{x}{a}\right|^{\nu} + \left|\frac{y}{b}\right|^{\nu} = 1 \qquad (1)$$

where a and b are the semi-axes in the x and y directions respectively and ν is a "squareness parameter" which controls the behavior of the loop radius of curvature. Fig. 1 illustrates the superquadric geometry for $\nu = 2$, 3, and 10 and an aspect ratio of $b/a = 2$. It is evident that the superquadric representation allows modeling of numerous different loop configurations through variation of the shape parameters a, b, and ν.

Fig. 2 presents the geometry for the two coupled superquadric antennas denoted as loops 1 and 2. The coordinates s_1 and s_2 are the lengths of the curves measured from each positive x-axis. The displacement vectors $\vec{r}_o$, $\vec{r}_1$, and $\vec{r}_2$ provide distances from the origin of coordinate system 1 as shown in the figure. The Eulerian angles α, β, and γ shown in Fig. 3 describe the rotations required to bring the coordinate frame of loop 2 onto that of loop 1.

Use of a parametric representation of the superquadric curve allows curved wire segments to be used in the moment method analysis. Observation has shown that a parametric form which provides mathematical simplicity and numerical stability assumes the form

$$\begin{aligned} x_p &= a_p \psi(\tau_p) \cos\tau_p \\ y_p &= b_p \psi(\tau_p) \sin\tau_p \end{aligned} \qquad (2)$$

where the subscript $p = 1$ or 2 denotes the loop number and $\psi(\tau_p)$ is given by

$$\psi(\tau_p) = \frac{1}{(|\sin\tau_p|^{\nu_p} + |\cos\tau_p|^{\nu_p})^{\frac{1}{\nu_p}}} \qquad (3)$$

for $0 \leq \tau_p \leq 2\pi$. The unit vector $\hat{s}_p$ in the directions of the length coordinate s_p may be obtained by properly differentiating (2).

Moment Method Solution

The parametric equation (2) for the superquadric curve is used in a coupled form of the EFIE for thin wires to represent the loop geometry and current. This EFIE is subsequently solved using a Galerkin type moment method employing piecewise sinusoidal subsectional basis and weighting functions. A thin wire assumption is used to overcome the difficulties associated with the singular integrand for coincident source and observation points. In this assumption, the observation point is on the wire axis and the source point is on the wire surface or on the wire axis depending upon whether the source and observation points are on the same loop or on different loops, respectively. Both delta gap and magnetic frill type source models are used as excitation schemes to allow investigation of different feeding scenarios. The current obtained from the moment method solution is subsequently used to obtain the antenna impedance

and radiation pattern for a given loop geometry. The radiation pattern due to the current on each loop is computed from the equation

$$\vec{E}_p(\theta,\phi) \approx \frac{-j\omega\mu_o}{4\pi}\frac{e^{-jk_o r}}{r}\int_{L_p} I_p(s'_p)\hat{s}'_p e^{jk_o\hat{r}\cdot\vec{r}_p{}'}ds'_p \tag{4}$$

where $I_p(s'_p)$ is the current on the pth loop.

Antenna Diversity

One of the objectives of this work is to determine the performance of the coupled superquadric loop antennas when used in a diversity scheme for a mobile communications system. In this application, we are interested in the use of space, angle, and polarization diversity to combat the effects of short-term or Rayleigh-type fading in a multipath environment. This type of fading occurs when multiple waves with random amplitudes and phases exist simultaneously to produce a spatial interference pattern, causing the received signal strength to fluctuate with antenna position. Antenna diversity scenarios in which multiple elements are used at the receiver are becoming more predominant as communications systems demand increased signal quality and reliability.

An important figure of merit for the performance of an antenna diversity configuration is the envelope correlation coefficient for the signals received by two different elements. In essence, this quantity provides a measure of the "similarity" of the two signals. For cases where the incident multipath field is assumed to arrive from the horizontal plane only it can be shown (see Vaughan and Andersen, 1) that the envelope correlation coefficient for two antennas may be computed from the equation

$$\rho_e = \frac{\left|\int_0^{2\pi}\vec{E}_1(\pi/2,\phi)\cdot\vec{E}_2^*(\pi/2,\phi)d\phi\right|^2}{\int_0^{2\pi}|\vec{E}_1(\pi/2,\phi)|^2d\phi\int_0^{2\pi}|\vec{E}_2(\pi/2,\phi)|^2d\phi} \tag{5}$$

where the $\vec{E}_p$ are obtained from (4) for the case of the coupled loop antennas.

EXAMPLES

Several interesting examples can be shown to illustrate the flexibility and usefulness of this computational tool. Unless otherwise indicated, a wire radius r_w is used such that

$$\Omega = 2\ln\frac{P}{r_w} = 10 \tag{6}$$

is satisfied, where P is the loop perimeter. Fig. 4 shows the current distribution for a single 0.25λ loop with $\nu = 10$ and aspect ratio $b/a = 1$. To obtain this current, a magnetic frill source model corresponding to a 50Ω feeding coaxial cable is used. Clearly, even for this small loop, the current distribution deviates significantly from the constant current often assumed in analysis. An illustration of the versatility of the mathematical construction appears in Fig. 5 where the directivity pattern for a folded dipole with $a = 0.00325\lambda$, $b = 0.25\lambda$, $r_w = 0.0005\lambda$, and $\nu = 50$ is shown. The antenna is fed using a 0.0065λ delta gap source. This is the familiar "doughnut" pattern associated with dipole radiators.

Fig. 6 provides the self and mutual admittance of two coupled circular loops as a function of the separation z_o for several values of circumference P. The dots in this figure correspond to data computed using a Fourier series approach for circular loop antennas by Iizuka et al (2). Clearly, excellent agreement exists between the two sets of data. Fig. 7 shows the variation in input impedance versus loop squareness ν for two concentric loops each 0.25λ at its respective operating frequency of f_o and $3f_o$. This plot compares the impedance of each loop when the other is short circuited to that of an isolated 0.25λ loop. Fig. 8 illustrates the variation of the envelope correlation coefficient as a function of antenna spacing for two parallel loops with $b/a = 1$ and $\nu = 10$ for different values of perimeter P. The dashed line in this figure provides the correlation coefficient for two antennas with isotropic radiation patterns in the horizontal plane. From this figure it is evident that envelope correlation coefficients below about 0.7 can be obtained using loops with spacings greater than approximately 0.15λ.

CONCLUSIONS

In this paper, a unified computational model has been presented which allows analysis of single or coupled superquadric loop antennas with arbitrary relative orientation and position. The analysis is performed using a Galerkin moment method with piecewise sinusoidal basis and testing functions. A parametrization for the superquadric geometry allows the use of curved wire segments in the analysis. The diversity performance of coupled loops has been discussed and demonstrated through an appropriate example. Several other examples have illustrated the usefulness and the versatility of the method.

Acknowledgements. This material is based upon work supported under a National Science Foundation Graduate Fellowship. It has also been partially funded by DARPA Contract #DAAB07-92-R-C977.

REFERENCES

1. Vaughan, R. and Andersen, J., 1987, IEEE Trans. Veh. Technol.,VT-36, 149–172

2. Iizuka, K., King, R., and Harrison, C., 1966, IEEE Trans. Antennas Propag.,AP-14, 440–450

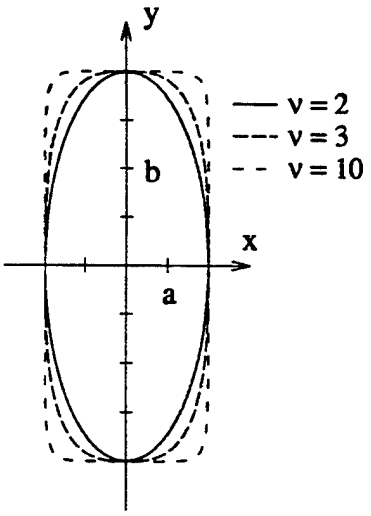


Figure 1: Superquadric geometry for $\nu = 2$, 3, and 10 with an aspect ratio of $b/a = 2$.

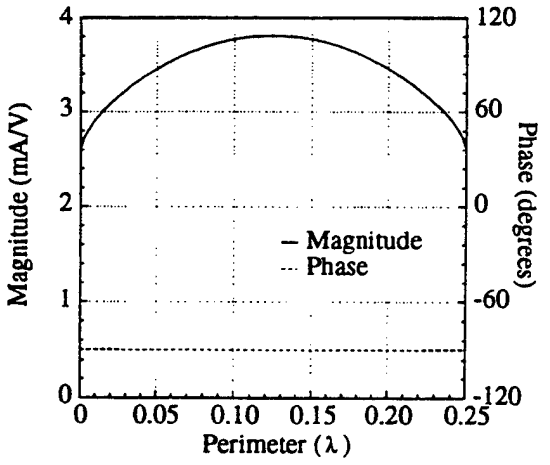


Figure 4: Current distribution (magnitude and phase) versus perimeter coordinate for a 0.25λ loop with $\nu = 10$ and $b/a = 1$.

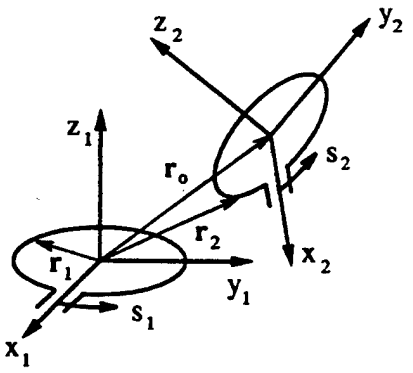


Figure 2: Geometry of two coupled loop antennas showing coordinates.

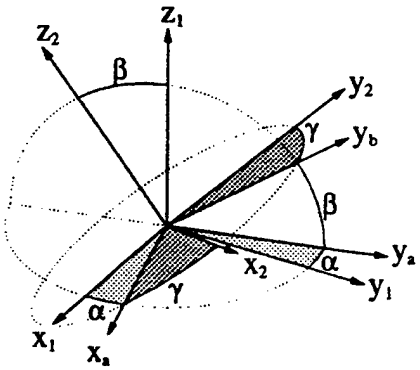


Figure 3: Eulerian angles α, β, and γ between coordinate systems 1 and 2.

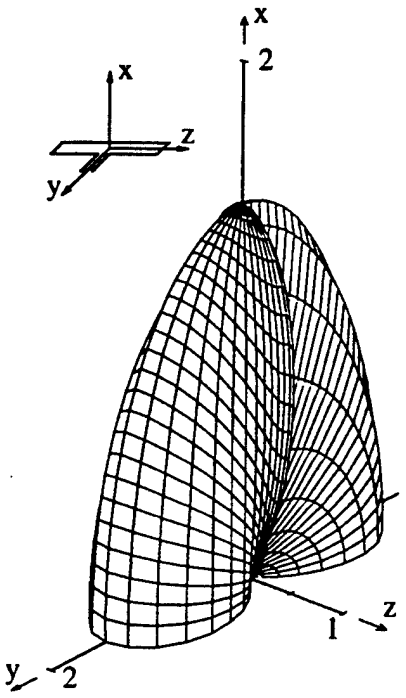


Figure 5: Directivity pattern in one hemisphere for a folded dipole with $r_w = 0.0005\lambda$, $a = 6.25r_w$, $b = 0.25\lambda$, and $\nu = 50$.

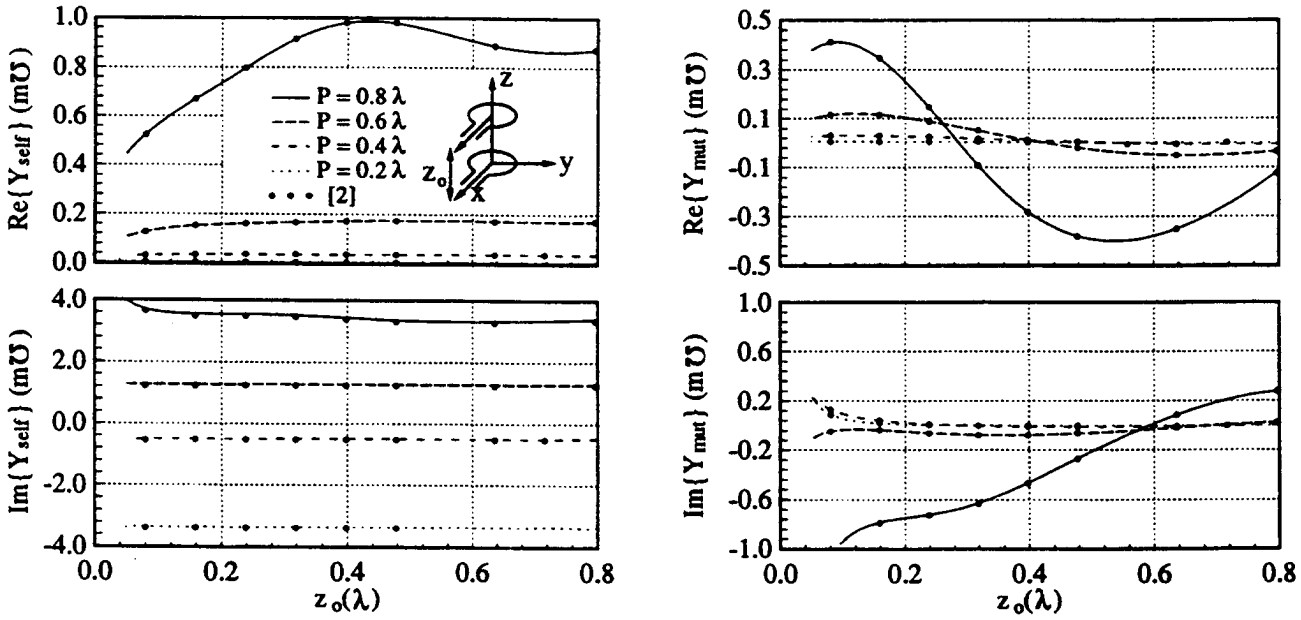


Figure 6: Self admittance versus separation distance for parallel circular loops with different loop circumferences P: (a) self admittance and (b) mutual admittance. The dots are computed data from Iizuka et al (2).

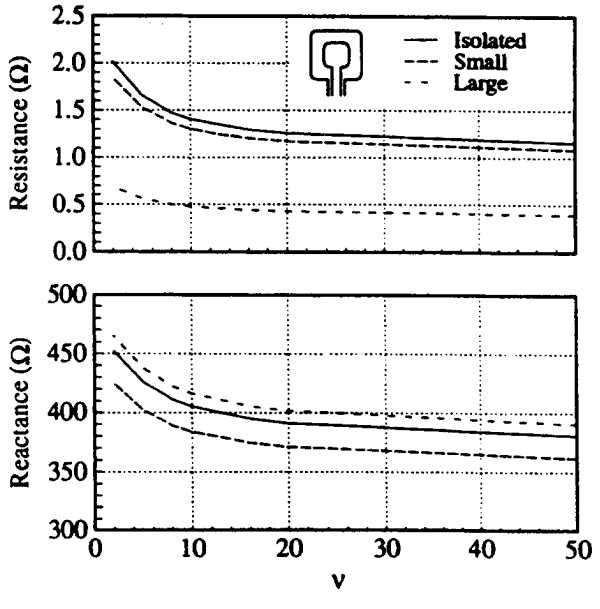


Figure 7: Input impedance versus squareness parameter ν for concentric 0.25λ loops compared with that of an isolated loop. The larger and smaller loops operate at f_o and $3f_o$ respectively.

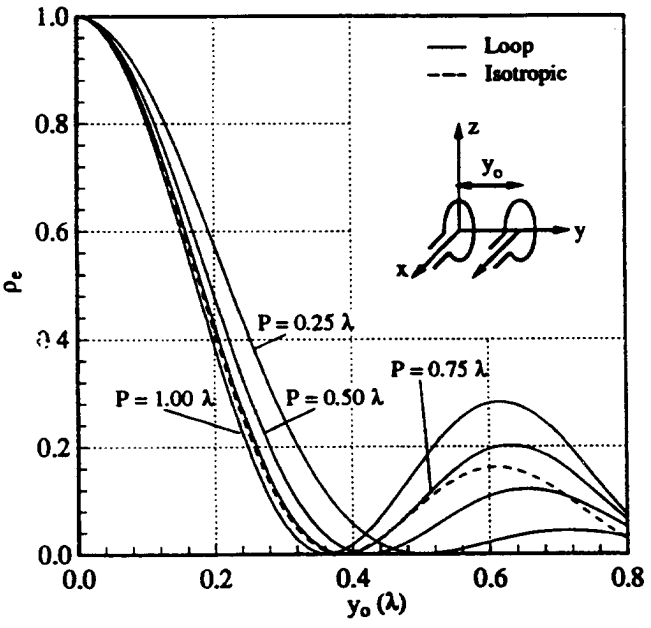


Figure 8: Envelope correlation coefficient versus separation distance for loops of perimeter P with $b/a = 1$ and $\nu = 10$ compared with that of two antennas with isotropic patterns in the horizontal plane.

Analysis of Isolated and Coupled Superquadric Loop Antennas

M.A. Jensen and Y. Rahmat-Samii

presented at the URSI National Science Meeting
Boulder CO
January 1993

B1-3
0920

ANALYSIS OF ISOLATED AND COUPLED SUPERQUADRIC LOOP ANTENNAS

M. A. Jensen and Y. Rahmat-Samii
Department of Electrical Engineering
University of California, Los Angeles
Los Angeles, CA, 90024-1594

The rapid expansion in the area of mobile communications has spurred the development of small, low-profile antennas suitable for implementation in portable equipment. In many devices such as pagers and hand-held transceiver units, the circular or non-circular loop antenna proves to be an effective radiator. Where packaging requirements limit the the possible loop shape and size, it is important to understand the effects of the antenna geometry on the input impedance and radiation characteristics. Furthermore, in applications where two loops are configured in an array or antenna diversity scheme, the effects of mutual coupling on the antenna impedance and diversity performance should be taken into account.

This work makes use of the mathematical construction of the superquadric curve to develop a very general formulation which allows modeling of a continuum of different loop geometries, including the well-known circular, square, and folded dipole configurations. The analysis capabilities allow investigation of the characteristics of isolated or coupled superquadric loops with different geometries, excitation, and relative orientation. The solution methodology employs a thin wire form of the electric field integral equation for arbitrarily shaped wires which is solved for the current distribution on the loop via a piecewise sinusoidal Galerkin moment method. The current obtained is subsequently used to find the input impedance, directivity, radiation pattern, and (in the case of coupled loops) self and mutual impedances for the antenna. In order to allow the use of curved wire segments, rather than the commonly employed piecewise linear segments, to construct the antenna geometry, a novel parametric form of the equation for the superquadric loop is developed. Both magnetic frill and delta gap source models are used to allow investigation of different possible feeding scenarios.

Application of the formulation to several superquadric loop configurations allows investigation of the variation of impedance and pattern characteristics with the antenna geometry. For the special case of the commonly-used circular loop antenna, the results are shown to agree favorably with previously computed results. The computations performed lead to some very useful curves which aid in the design of practical superquadric loop radiators. The results provided include the effects of antenna position, excitation, and geometry on the self and mutual impedance of two coupled loops. The current distribution on the coupled loops allows determination of the envelope correlation coefficients for signals received by antennas arranged in a diversity configuration. This computation provides a figure of merit of the system's ability to combat the effects of multipath fading in a typical mobile communications environment.